


STARS, VOTES, AND BADGES: HOW ONLINE BADGES AFFECT HOTEL REVIEWERS

Markus Schuckert 
Xianwei Liu
Rob Law

ABSTRACT. As online reviews, ratings, and helpful votes in e-commerce platforms become industry-centric, this study examines how virtual badges affect the online behavior of reviewers and readers based on status-seeking theory in an online environment. This study finds that online readers prefer reviews with low ratings to those with high ratings; however, reviewers with high-level badges tend to post moderate ratings and avoid extreme ratings. Casting doubt on the efficiency of current online incentive systems, user-generated content websites may establish a qualitative index to measure the quality of customer reviews from a long-term perspective to encourage reviewers to continue creating high-quality contents.

KEYWORDS. Experience, motivation, reputation, online badges

1. INTRODUCTION

The number of online reviews is currently increasing. For example, TripAdvisor provides over 40 million traveler reviews and ranked lists of over 125,000 visitor attractions, 450,000 hotels, and 600,000 restaurants (Jeacle & Carter, 2011). Thus, finding an answer to the question “What is the motivation behind posting online reviews?” is valuable not only from an industry viewpoint but also from an academic one.

Yoo and Gretzel (2011) found that women are more motivated by their desire to help a company, whereas men are motivated by their desire to prevent others from falling into traps. Similarly, Ögüta and Cezara (2012) found that higher ratings and lower prices increase the propensity to post reviews. Complaints about bad experiences also motivate customers to post reviews, and are likely to result in bad reviews. Sparks and Browning (2010) stated that customers who experience service failure tend to spread negative word-of-mouth (WOM)

Markus Schuckert, Assistant Professor, School of Hotel and Tourism Management, The Hong Kong Polytechnic University, 17 Science Museum Road, TST-East, Kowloon, Hong Kong SAR (E-mail: markus.schuckert@polyu.edu.hk).

Xianwei Liu, PhD Candidate, Harbin Institute of Technology, School of Management, 92 Xidazhi Street, Harbin, 150001, China (E-mail: liuxianwei@126.com).

Rob Law, Professor, School of Hotel and Tourism Management, The Hong Kong Polytechnic University, 17 Science Museum Road, TST-East, Kowloon, Hong Kong SAR (E-mail: rob.law@polyu.edu.hk).

Address correspondence to: Markus Schuckert, School of Hotel & Tourism Management, The Hong Kong Polytechnic University, 17 Science Museum Road, TST-East, Kowloon, Hong Kong SAR (E-mail: markus.schuckert@polyu.edu.hk).

The authors acknowledge the financial support of the Hong Kong Polytechnic University [grant number A-PM08].

and have the potential to influence the reputation of a brand or firm. The motivation for posting negative WOM can be to take revenge or warn others (Wetzer, Zeelenberg, & Pieters, 2007). Restaurant service employees who create a good experience can also trigger positive WOM, which is motivated by the desire to support the employees and the restaurant, or simply to express personal positive feelings (Jeong & Jang, 2011).

The research stream is heading in the direction of external factors to discover the triggers and motivations behind such activities. Empirical findings showed that online travel reviewers providing feedback are motivated by their desire to support the service provider and to push for improvements in service quality. They also want to complain about bad experiences and express their feelings as well as concern for future consumers and their experiences. However, existing literature does not include studies that focus on how internal factors can motivate travelers to post online reviews; an internal factor can, for example, take the form of a reward system on the platforms where reviewers can post reviews or ratings.

Online reviewers can also be motivated to post more reviews by earning higher-level badges that are offered by the reward systems of websites. Reviewers who create more posts can thus obtain a more prestigious online badge. Therefore, frequent posting offers the opportunity to ascend the reputation ladder and demonstrate that one's contributions are based on experience and expertise. A high-level badge signals that the reviewer is authoritative, senior, and experienced (Mkono, 2012; Yoo & Gretzel, 2010). Users of e-commerce platforms mainly comprise reviewers and readers. Reviewers are contributors who write numerous reviews that provide potential customers (readers of these reviews) with some or all the information they need to make "the right" decisions (Mauri & Minazzi, 2013). Nearly all e-commerce platforms offer these functions, systems, and incentive hierarchies to encourage their members to post more reviews and, in turn, benefit their business (Sparks & Browning, 2011).

These complex systems of online rewards, such as badges, aim to encourage users to contribute more content to increase their status or reach different levels. However, the initial meaning of giving a review changes in terms of contributing to a system to increase personal value (Zhang, Ye, Law, & Li, 2010). Thus, an important task is to ask whether and how these reward systems change the behavior of contributors and readers. We also need to understand the effect of this system on customers and its wider implications for the industry. On this basis, we examine whether different online badge levels affect the online behavior of both reviewers and readers. This question itself implies that online badge levels may have an effect on these groups, which leads to a question about the difference and diversity of actions. That is, do reviewers with high badges act differently from those with less prestigious badges when providing ratings? The findings of this study can also offer insights into the answer to the question with regard to the kind of online reviews that readers prefer.

Few studies have focused on value creation or the relationship between website badges and review ratings, with most of the current literature focusing on the helpfulness of online reviews. The findings of such work are generally focused on which reviews are considered more helpful and likely to attract more votes (Black & Kelley, 2009; Lee, Law, & Murphy, 2011; Zhang & Tran, 2010), and what kind of review is considered more authentic (Weimann, Tustin, Vuuren, & Joubert, 2007). In our view, understanding the effect of these reward systems on behavior and identifying the preferences of users are important. Reviewer badges are more than just a status symbol. In the online environment, reviewers may change their behavior on the website as their badge level rises. In this research, we focus on the effect of badges, which represent travel experience and serve as status symbols, on the ratings and value creation of reviewers. Specifically, we examine whether reviewers with high-level badges post moderate ratings, avoid extreme ratings, or create more helpful reviews.

In the hospitality industry, online ratings are important, and the number of websites that

provide such ratings has recently increased dramatically. This situation is due to the special characteristic (namely the intangibility) of services provided by the industry, such as offering hotel rooms. These services cannot be tested or used in advance and then returned if the customer is not satisfied. In this study, we use the reward system of TripAdvisor as a source of data set to test our hypotheses and answer the question: Do the online incentive hierarchies work? Thus, this research differs from other current studies. For example, as stated before, empirical findings have shown that online reviewers providing feedback are motivated by a desire to support the service provider and to push for improvements in service quality, as well as a concern for future consumers and their experiences (Jeong & Jang, 2011; Yoo & Gretzel, 2008, 2011). Sparks and Browning (2010) found that customers who experience service failure are motivated to spread negative WOM and have the potential to influence the reputation of a brand or firm.

In the following, we provide an overview of recent literature and other relevant findings and develop our hypothesis based on these insights. We then outline our research design and explain the sampling and data collection procedures. Next, we present the results and discuss the findings, including an analysis of the implications for users, industry, and researchers. The study concludes with a critical view of its limitations and opportunities for future research.

2. BACKGROUND AND HYPOTHESES

2.1. Status-Seeking Theory and Online Behavior

The underlying motivation for posting reviews is a psychological incentive for individuals to disseminate positive WOM or satisfaction, that is, to gain social approval or self-approval (Chen, Fay, & Wang, 2011; Fehr & Falk, 2002), to signal their expertise and experience, or to gain social status (Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004).

Relating to online environment and based on the work of Berger, Cohen, and Zelditch (1972) and Thye (2000), Lampel and Bhalla (2007, p. 437) perceive status seeking as activities that are “designed to improve an actor’s standing in a group, and is therefore judged by the degree to which associated activities result in increasing prestige, honor, or deference”. Status seeking can be externally motivated, specifically, seeking economic and social advantage, or internally motivated for psychological and emotional reasons (Perretti & Negro, 2006). Status seeking is linked to self-image (Köszegi, 2006), gaining public recognition (Moldovanu, Sela, & Shi, 2007; Rustichini, 2008), or outperforming others (Dohmen, Falk, Fliessbach, Chowdhury & Sheremeta, 2012; Dohmen, Falk, Fliessbach, Sunde, & Weber, 2011).

Users who post reviews are more likely to be status seeking and to engage in conspicuous consumption than users who just consume the content (Cowan, Cowan, & Swann, 2004). Chen et al. (2011) found that demonstrating expertise, experience, and social status is an important part of virtual communities. Hence, status seeking, reputation gaining, and conspicuous consuming are phenomena that are observed offline and online.

2.2. Reviewer Badges and Rating Behavior

Many types of website employ a badge system, such as e-commerce platforms (eBay, Amazon, and Taobao), online travel agents (Ctrip, TripAdvisor, and Agoda), and online question-and-answer sites (such as Yahoo! Answers, Answerbag, and StackExchange). Antin and Churchill (2011) found that badges provide information from which a reputation assessment can be made. Badges display a user’s choice, expertise, and past experience (Kollock, 1999). Furthermore, badges show engagement, experience, and expertise, as well as signaling trustworthiness within the virtual community. Relating to status, badges can be motivating as status symbols (Antin & Churchill, 2011) or, in Lampel and Bhalla’s

(2007, p. 437) understanding, as an “ego reward” or an emotional good that a user “can accumulate as a result of acquired status”. Badges provide personal affirmation by serving as reminders of past achievements much like trophies on a mantelpiece, marking milestones, and providing evidence of past success (Antin & Churchill, 2011). Badges are a symbol of reputation, as the online community may view them as an indicator of expertise and experience, which may then influence their future behavior (Antin & Churchill, 2011; Meng, Webster, & Butler, 2013).

The badge level represents a participant’s status in the online travel community. An individual may therefore seek to earn higher badge levels in order to gain public recognition, which results in certain contribution behaviors (Thompson, 2005). The badge systems of different websites work in similar ways. For example, participants in question-and-answer sites can earn badges when their answers are accepted or when they accomplish specific tasks. Reviewers who contribute to e-commerce platforms and online travel agencies (OTAs) can earn badges, tags, or medals by posting more reviews (Butler & Wang, 2012). The purpose of badge systems is to engage members and encourage greater participation (Li, Huang, & Cavusoglu, 2012). In addition, based on theories of status seeking and self-presentation, providing opinion, advice, and experiences in online platforms is heavily influenced by the desire for status as an intrinsic motivation and external non-monetary reward (Lampel & Bhalla, 2007). Non-monetary-based mechanisms, such as titles, stars, votes, and badges, seem to be more effective than monetary-based mechanisms (Li et al., 2012; Shah, Oh, & Oh, 2008).

TripAdvisor categorizes its badges into five levels based on reviewer activity, as shown in Figure 1. Users who have posted fewer than

three times are not recognized. Three to five reviews are required to gain the status of “reviewer”. The thresholds for subsequent categories are exponential. After six reviews, one earns the status of “senior reviewer”; with 11 reviews, one becomes a “contributor”; with more than 20 reviews, one becomes a “senior contributor”; and with more than 49 contributions, one becomes a “gold member” or “top contributor”. A rising status is also denoted by assigning a symbolically more “valuable” graphic to each stage, with a light green star at the bottom of the scale and a solid gold star at the top.

A reviewer posts a comment and rating by using his or her account and username. Every subsequent review posted using the same account will be credited to that user’s reviewing statistics. If the number of reviews crosses one of the predefined thresholds, then a new badge status is assigned to that account. Reviewers can earn a higher badge by posting more reviews. Thus, the badge level is an indication of a reviewer’s travel experience (Chen & Huang, 2013). The number of reviews is also an indicator of travel activity and frequency, because visiting more hotels means that one can contribute more reviews. The rating platforms usually allow only customers with an actual booking (and supposedly a real stay) to file a review (Vermeulen & Seegers, 2009). An increasing number of booking engines and online booking platforms are now linked to TripAdvisor. Customer data, booking platforms, and mail robots are integrated systems, which means that customers can be asked automatically to rate their stay after a trip has been completed (Bronner & De Hoog, 2011). Numerical ratings for online reviews typically range from one star to five stars (Zhang, Zhang, Wang, Law, & Li, 2013). A very low rating (one star) indicates an extremely negative view of the product or

FIGURE 1. Example for Badge Levels



Source. TripAdvisor

service, a very high rating (five stars) reflects an extremely positive view, and a three- or four-star rating reflects a moderate assessment (Mudambi & Schuff, 2010). Some star ratings reflect extreme attitudes toward a product or service, as indicated by a deviation from the midpoint of the attitude scale (Krosnick, Boninger, Chuang, Berent, & Carnot, 1993). For example, a three-star review can reflect a moderate view, which can indicate indifference, but it can also comprise a series of positive and negative comments that cancel each other out, which indicates ambivalence.

The kind of reviewer who likes to give extreme ratings is generally unknown. In this study, we define extreme ratings as one-, two-, and five-star ratings. A question that should be addressed is whether the badge levels and the travel experience that these ratings represent are considered important by reviewers. For example, experienced and inexperienced travelers differ in many aspects (Yoo & Gretzel, 2008). Experienced travelers know what to expect when visiting a hotel or restaurant, and they understand what a place can or cannot deliver, or what is within acceptable limits and what is not (Jeong & Jang, 2011). Furthermore, experienced travelers may be more patient than inexperienced ones and are also likely to be more professional when giving ratings and posting reviews (Shanteau, Weiss, Thomas, & Pounds, 2002; Willemsen, Neijens, & Bronner, 2012). Inexperienced travelers tend to be overcritical because they do not understand the industry well, have not formed a realistic set of expectations, and do not know who they should hold accountable if something goes wrong (Black & Kelley, 2009). Therefore, inexperienced and experienced reviewers act differently when they rate services or products. Experienced reviewers with a higher status are more objective and unbiased, whereas inexperienced reviewers are harder to please and seek status (Kim, Mattila, & Baloglu, 2011). This finding can be related to the findings of Kemper (1991) or Merton (1968), where high-status and experience-rich individuals tend to be more (positively) balanced than (aggressive) status seekers. According to this analysis, we propose the first hypothesis (H).

H1: Reviewers with high-level badges will be more likely to give moderate than extreme ratings.

2.3. *Helpful Votes and Value Creation*

As noted earlier, we know that online platforms such as Amazon, eLong, and Qunar allow readers to give helpful votes to contributions or comments made by reviewers. This kind of feedback assures the quality and credibility of posts, as well as constituting a second-level source of support for potential customers (readers are more likely to rely on the customer reviews that are marked with more helpful votes). Readers rate reviews as helpful when they consider the provided information to have been useful to them. This situation can take place before a trip when customers are looking for potential accommodation. The readers vote by clicking on an icon. The total number of clicks represents the value of the review, which means that a high number of votes can indicate greater helpfulness (Lee et al., 2011).

This finding raises the question of what makes an online review helpful and valuable. Pavlou and Dimoka (2006) show that extreme ratings are more influential than moderate ones on eBay. Forman, Ghose, and Wiesenfeld (2008) similarly show that, for books, moderate reviews are considered less helpful than extreme ones. However, Baek, Ahn, and Choi (2012) found the opposite to be the case in their study of Amazon reviews, which showed that a greater difference between review star rating and product average rating corresponds to a lower helpfulness score. They report that higher star and product average ratings result in fewer helpfulness votes. Sen and Lerman (2007) explain this apparent contradiction by suggesting that the degree of helpfulness depends on product type regardless of the positive or negative tone of the actual review. Mudambi and Schuff (2010) show that product type controls the effect of an extreme review on the helpfulness score, given that the readers who view the products have different

information needs. Therefore, not all low ratings or negative reviews will meet the readers' requirements and produce helpfulness votes. Ultimately, the helpfulness of a review is related to product type.

Studies also show that reviewer attribution can affect the perceived helpfulness of reviews. According to Forman et al. (2008), reviews that disclose identity-descriptive information about the reviewer are rated as more helpful than anonymous ones. Meanwhile, Lee et al. (2011) indicate that the travel frequency and ratings of reviewers can also affect the helpfulness score that they give to their reviews. In this manner, the reviewer badge level, which represents experience, authority, and status, is an important reviewer attribute. However, as mentioned earlier, research on the effect of reviewer badge level on value creation remains lacking. Based on the results of these studies, we ask whether reviewers with higher-level badges generate more value and obtain a higher proportion of helpfulness votes than reviewers with lower-level badges.

We infer that a reviewer with a high-level badge will contribute less than one with a lower status for several reasons. Firstly, reviewers who already have a high-level badge will be less motivated to contribute a review. When someone initially joins such websites, he or she may aim to collect badges to obtain more recognition from readers, in a manner similar to playing online games. However, once a reviewer reaches the highest badge level, he or she may no longer be motivated to participate. The quality or value of their reviews may decline, which results in decreasing helpfulness votes. Secondly, badges do not only represent experience, but also status on the site where they appear. After a reviewer has achieved this status, he or she may become less competitive than other users (Antin & Churchill, 2011). Based on this analysis, we propose our second hypothesis.

H2: Reviewers with high-level badges will post fewer helpful reviews and receive a lower proportion of helpful votes than reviewers with low-level badges.

3. METHODOLOGY

3.1. Data and Variables

TripAdvisor is the best-known platform for online reviews in the hospitality industry and is increasingly considered as a metric for the quality of hotels and other tourism-related services. In early 2013, TripAdvisor had over 200 million monthly unique users, according to Google Analytics (Lee et al., 2011). However, we did not include all users in our analysis because of the considerable amount of time and effort required to process such data. Instead, we selected reviewers from different regions and hotel classes to ensure that a suitably unbiased sample was chosen (Ye, Law, Gu, & Chen, 2011). We chose hotels from all classifications ranging from one star to five star in Hong Kong, listed on TripAdvisor, collected information on all reviewers associated with these hotels, and omitted duplicates. This approach ensured both the size and validity of the sample. Hong Kong is an international city that attracts travelers from all over the world; hence, the reviewer sample is diverse. Using a crawler, selected web page related data were downloaded that show reviewer badges, review ratings, and helpful votes (Ye, Zhang, & Law, 2009). In addition, another Java-based program was developed to parse HTML- and XML-web pages into our database. Data collection was conducted in August 2013. The crawler only requires several seconds to collect data on each reviewer. Thus, the status of each reviewer will not change during the data collection period. The actual benefits (monetary rewards or coupons) that users obtain from various platforms differ. For example, Qunar, Ctrip, and eLong offer points or coupons that can be used by people who post reviews, whereas Booking.com provides discounts. Platforms like TripAdvisor, however, do not offer any kind of tangible rewards (monetary rewards or coupons). TripAdvisor for example, only offers badges to reviewers. This practice is the main reason why TripAdvisor was chosen for analysis in this study because the effect of an actual monetary

or quasi-monetary benefit on reviewer behavior can be excluded.

Each review has a corresponding number of helpful votes. If readers consider a review to be helpful, then they can vote for it; if a review is not considered to be helpful by any reader, then it will receive no vote. We collected the number of helpful votes received and the number of ratings posted by each reviewer, including excellent, very good, average, poor, and terrible. As discussed earlier, the extremeness and helpfulness of a review can be measured by its rating and helpful votes. However, the review content itself can also be a factor. Thus, at least two approaches are available to measure the extremeness and helpfulness of a review, namely the score and the content. Text mining (sentiment analysis) of review content can classify online reviews into positive or negative, or helpful or unhelpful, and has been a popular research topic over the past few years (Ye, Zhang, et al., 2009; Zhang, Ye, Zhang, & Li, 2011). However, the accuracy rate is only approximately 70% and the biggest challenge is conducting such an analysis in different languages, which will further reduce accuracy rate. Consequently, this study used ratings rather than content to measure review extremeness and helpfulness. The variables in our research are defined in Table 1.

3.2. Research Design

Three steps are employed in our research. Firstly, we transformed TripAdvisor badge levels into their equivalent numerical values (from 1 to 5) to conduct the empirical study (Table 2). We excluded the 0 level. According to TripAdvisor (see Figure 1), a reviewer receives no badge if he or she posts fewer than three reviews. For example, if a reviewer posted 0/1/2 reviews, then the corresponding badge level would be 0. These reviewers were excluded because an extremely small number of reviews is inappropriate or insufficient to calculate variable ratio (excellent, very good, average, poor, terrible). If a reviewer has only posted one review, then the “excellent” ratio will be 1/1 or 100%, whereas other ratios such as “very good” and “average” will be 0. This situation will lead to bias because ratio or the probability of each kind of score (from excellent to terrible) should not be calculated using only a few reviews (that is fewer than 3). Given that we cannot measure the satisfaction degree of a hotel by using only one or two ratings, we exclude hotels with fewer than 10 ratings to ensure that bias will not be introduced (Liu, Schuckert, & Law, 2014; Schuckert, Liu, & Law, 2014).

Secondly, we computed every variable at the reviewer level rather than at the review level.

TABLE 1. Variables Used for Analysis

Variables	Description
Badges	Badge levels, based on the badges given by TripAdvisor (see Table 2)
Hratio	The number of helpfulness votes/The total number of ratings for each reviewer
Excellent	Excellent ratio: The number of “excellent” ratings/The total number of ratings for each reviewer
Very Good	Very good ratio: The number of “very good” ratings/The total number of ratings for each reviewer
Average	Average ratio: The number of “average” ratings/The total number of ratings for each reviewer
Poor	Poor ratio: The number of “poor” ratings/The total number of ratings for each reviewer
Terrible	Terrible ratio: The number of “terrible” ratings/The total number of ratings for each reviewer

TABLE 2. Numerical Definitions of Badge Levels

Badges	Reviewer	Senior reviewer	Contributor	Senior contributor	Top contributor
Levels	1	2	3	4	5

The reviewers for TripAdvisor can post three kinds of online review/rating: hotels, restaurants, and attractions. TripAdvisor does not differentiate between a review of a hotel, a restaurant, or an attraction when calculating the number of reviews a reviewer has posted. Thus, the effect of badges on reviewer behavior should be consistent with the badges arranged by TripAdvisor. In this study, we collected data on badges based on the total number of reviews rather than on the number of hotel reviews. When we calculate the extremeness ratio (Eratio) and the helpfulness ratio (Hratio), we also used all the reviews or ratings posted by a reviewer. For example, the Hratio was computed as follows. If a reviewer posted 10 reviews in total and gained three “helpful” votes, the Hratio of this reviewer was computed as $3/10 = 0.3$, and so on for other variables. Thirdly, we computed the ratio of each kind of rating (excellent, very good, average, poor, and terrible) for each reviewer. We then ran a correlation analysis between reviewer badge levels and each of the rating level to explore our hypotheses.

4. RESULTS AND DISCUSSION

4.1. Descriptive Statistics

Data for 43,764 reviewers were selected and deemed valid after omitting those with missing values (Table 3). In addition, 10,451 reviewers (who posted fewer than three reviews) have no badge. The total number of reviews posted by all reviewers is 1,181,935. Badge levels ranged from 1 to 5, with an average value of 3. By contrast, the range of the Hratio was significantly wide, that is, from 0 to 10, with an average value of 1. This finding indicates that the average reviewer receives only one “helpful” vote (for all their reviews, rather than just one). A reviewer can give ratings from 1 to 5. According to the descriptive statistics, most reviewers tended to give “excellent” and “very good” ratings, with an average of 76.3%. This result suggests that most travelers were satisfied with the hotels or restaurants that they had visited.

TABLE 3. Descriptive Statistics of Variables

	Minimum	Maximum	Mean	Standard	N
Badges	1.00	5.00	3.031	1.317	43,764
Hratio	0.00	10.00	1.002	1.088	43,764
Excellent	0.00	1.00	0.399	0.233	43,764
Very good	0.00	1.00	0.364	0.194	43,764
Average	0.00	1.00	0.155	0.132	43,764
Poor	0.00	1.00	0.052	0.079	43,764
Terrible	0.00	1.00	0.030	0.069	43,764

Among the reviewers, 15.5% regarded their experience as “just okay”, whereas only a small proportion (8.2%) rated their satisfaction as “poor” or “terrible”.

4.2. Rating Behavior

Firstly, we analyzed the relationship between reviewer badge levels and ratings (1–5). A significant negative relationship exists between badge level and “excellent”/“poor”/“terrible” ratings, whereas a positive relationship exists between badge level and “very good”/“average” ratings (Table 4). This finding indicates that reviewers with high-level badges are more cautious than reviewers with low-level badges, and thus are less likely to give extreme ratings such as 1, 2, or 5. In most cases, they tend to give relatively moderate scores such as 3 and 4.

Secondly, as defined in H1, we conducted a correlation analysis between reviewer badge level and review extremeness. We measured the latter in two ways (scores of 1 and 5 and scores of 1, 2, and 5) because a rating of 2 is considered as extremely negative, and only a few reviewers give ratings of 1 or 2. Badge level is significantly negatively correlated with extreme ratings regardless of how the latter are defined (Table 5). In this analysis, an extreme review score is computed for each reviewer and the measure of extremeness is equal to the proportions of 1, 2, and 5, or the scores of 1 and 5.

Notably, reviewers with high-level badges tend to post reviews and ratings in a cautious and responsible manner. We suggest two possible reasons why such reviewers are more responsible and objective than reviewers with a lower status. Firstly, they are aware that their

TABLE 4. Correlation Analysis

	Badge	Hratio	Excellent	Very good	Average	Poor	Terrible
Badge	1.000						
Hratio	-0.160**	1.000					
Excellent	-0.126**	-0.006	1.000				
Very good	0.113**	-0.059**	-0.692**	1.000			
Average	0.134**	-0.011*	-0.539**	-0.030**	1.000		
Poor	-0.031**	0.093**	-0.252**	-0.165**	0.038**	1.000	
Terrible	-0.114**	0.102**	-0.106**	-0.228**	-0.058**	0.095**	1.000

Notes. Hratio: helpfulness ratio; **Correlation is significant at 0.01 level (2-tailed); *correlation is significant at 0.05 level (2-tailed).

TABLE 5. Badge Level and Extreme Review Score Correlation Analysis

Extreme review score	Reviewer badge level
(Scores 1, 2, and 5)	-0.171**
(Scores 1 and 5)	-0.158**

Notes. **Correlation is significant at 0.01 level (2-tailed).

reviews or ratings will influence more travelers; hence, they rate carefully and avoid any kind of extreme reaction (too positive or too negative). Secondly, they are more experienced and have traveled more, and thus they can make better comparisons and offer more objective ratings. Therefore, H1 is supported.

4.3. Value Creation

We also examined whether reviewers with high-level badges produce more helpful and valuable information and whether the quality or value of a review declines as a reviewer posts more reviews. We therefore conducted a correlation analysis between reviewer badge level and Hratio to explore these issues.

A significant negative relationship exists between Hratio and reviewer badge level (coefficient: -0.160; significance level: 0.01) (see Table 4). This finding indicates that the review quality of reviewers with high-level badges, who are highly productive (they post more reviews) decreases, and thus these reviews receive fewer “helpful” votes. That is, although such reviewers post more, the value of their

contribution to readers declines. Hence, H2 is supported.

This phenomenon has many potential explanations. The most direct explanation is that readers prefer reviews with low ratings, but reviewers with high-level badges tend to post more moderate reviews and avoid very high or low ratings. A significant positive relationship exists between Hratio and the scores of 1 and 2, whereas a negative relationship exists between Hratio and the scores of 3 and 4 (see Table 4, second row). By contrast, reviews with low ratings (1 and 2) receive more “helpful” votes. This result indicates that readers find bad reviews (namely those with low ratings) more helpful, and thus vote for them accordingly. In addition, reviews may also be read to assist in making travel decisions, and most potential customers try to avoid risks by focusing on bad reviews. When readers find several bad reviews, they may postpone their travel plans or consider another hotel. Hence, excellent (insignificant) reviews may be completely ignored (Lee, Park, & Han, 2008).

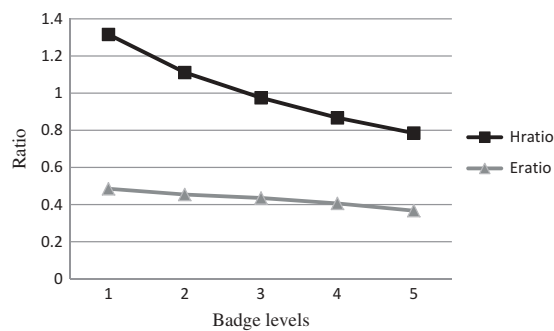
Another possible explanation is that the quality of reviews is reduced as badge status rises. In order to explore the relationship between reviewer badge level and extreme review scores and value creation, we computed Eratio and Hratio according to reviewer badge levels (1–5). Eratio is defined as the proportion of 1 and 5 scores respectively of each reviewer, with the average value computed according to reviewer badge levels; the same procedure is applied to Hratio (Table 6). As shown in Figure 2, Eratio is reduced as badge status rises, which indicates

TABLE 6. Badge Levels, Hratio, and Eratio

Badge levels	Hratio	Eratio
1	1.316	0.485
2	1.111	0.454
3	0.975	0.435
4	0.867	0.406
5	0.785	0.367

Notes. Hratio: helpfulness ratio; Eratio: extremeness ratio.

FIGURE 2. Helpfulness ratio (Hratio), Extremeness ratio (Eratio), and Badge Levels



that reviewers with a higher status are less likely to give extreme ratings. Hratio is also reduced when badge status rises, which suggests that reviewers with high-level badges produce less valuable reviews. Such reviewers may care less about quality once they have reached a relatively high status compared with when they were still working to gain status (through badge level). Current e-commerce platforms only offer the absolute quantity of “helpful” votes. Reviewers with high-level badges have already gained numerous votes and do not need to work for one or two more. However, if platforms supply an index that reflects the contribution rate of a reviewer, which is a relative value, the situation may considerably change and reviewers may be inspired to post more valuable content in the long run.

5. CONCLUSION

To our knowledge, this study is one of the first empirical ones to focus on the effects of

travel experience and online badges on reviewer behavior (review extremeness and value creation), which can help to improve our understanding of online reviews. It also contributes to status-seeking theory in an online environment. Online reviewers on OTA platforms can be motivated by badge levels to increase their status in their network life. This status-seeking behavior has changed the rating behavior and value creation of online reviewers. Two important insights can be gained from our empirical findings. Firstly, reviewers with high-level badges dislike giving extreme ratings and tend to give relatively moderate ratings. Secondly, although such reviewers produce more content, the quality of their reviews decreases and is considered to contribute less to readers. In addition, readers find reviews with low ratings to be more helpful than those with high ratings, and thus are more likely to give “helpful” votes to the former within the context of the travel and hotel industries.

Based on these findings, we suggest that applications or platforms such as TripAdvisor and other OTAs must establish a new index to measure the quality of reviews, given that existing indices focus only on quantity (even “helpful” votes, which represent review quality, are measured by an absolute index rather than a ratio). A quality index may encourage reviewers to continue creating valuable contents, which will benefit the sustainable development of rating platforms. In addition, social hierarchy also works in the online environment. That is to say, high-status reviewers create more careful and responsible reports, although readers are more anxious to learn about the negative aspects of a product or service.

Our research has several limitations. Firstly, we did not analyze the personal characteristics of the reviewers, even though background variables can significantly affect online behavior. Gender, age, and other personal information were missing from our data set because they are regarded as private by most TripAdvisor members. Future research may consider the personal characteristics of reviewers when exploring their behavior. Secondly, the correlation coefficients generated in our analysis were insufficient, and thus the same application can

be used on different platforms such as Ctrip and eLong to further verify the proposed hypotheses. Thirdly, different platforms have various restrictions on posting reviews. For example, Agoda, HotelClub.com, Ctrip, and eLong only allow customers who have made bookings to post reviews and ratings. Other platforms such as TripAdvisor and Qunar allow users to post reviews by asking them to declare that the review is based on actual experience. This condition can influence the robustness of our results. When studying online reviews, the practice of assuming that real reviews exist (reviews posted by customers with actual bookings) is more relaxed. The number of fake reviews may be minimal, and thus can only slightly influence the findings of this study. Fourthly, online ratings are a necessary but insufficient proxy to measure the opinion of reviewers. Therefore, future studies could combine content and rating to measure the opinion of customers or to explore consistency or heterogeneity between them. Lastly, given that a proportion of wrong clicks on the “helpful” vote button occur, a bug is possibly present in TripAdvisor and the validity of the number of “helpful” votes is reduced. Users should therefore be reminded that they are casting their vote when clicking this button.

In this study, we only collected a cross-section data set of each reviewer and explored various online behaviors among reviewers with different badge levels. A future research direction will be using a panel data set to search for dynamic changes in the rating behavior of reviewers.

ORCID

Markus Schuckert  <http://orcid.org/0000-0003-1912-8672>

REFERENCES

- Antin, J., & Churchill, E. F. (2011, May 7–12). *Badges in social media: A social psychological perspective*, CHI 2011 Gamification Workshop Proceedings, Vancouver, BC, pp. 1–4.
- 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. doi:10.2753/JEC1086-4415170204
- Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, 37(3), 241–255.
- Black, H. G., & Kelley, S. W. (2009). A storytelling perspective on online customer reviews reporting service failure and recovery. *Journal of Travel & Tourism Marketing*, 26(2), 169–179. doi:10.1080/10548400902864768
- Bronner, F., & De Hoog, R. (2011). Vacationers and eWOM: Who posts, and why, where, and what? *Journal of Travel Research*, 50(1), 15–26. doi:10.1177/0047287509355324
- Butler, B. S., & Wang, X. (2012). The cross-purposes of cross-posting: Boundary reshaping behavior in online discussion communities. *Information Systems Research*, 23(3/II), 993–1010. doi:10.1287/isre.1110.0378
- Chen, H.-N., & Huang, C.-Y. (2013). An investigation into online reviewers' behavior. *European Journal of Marketing*, 47(10), 1758–1773. doi:10.1108/EJM-11-2011-0625
- Chen, Y., Fay, S., & Wang, Q. (2011). The role of marketing in social media: How online consumer reviews evolve. *Journal of Interactive Marketing*, 25(2), 85–94. doi:10.1016/j.intmar.2011.01.003
- Chowdhury, S. M., & Sheremeta, R. M. (2012). Strategically equivalent contests. Working Paper 10-06, Economic Science Institute, Chapman University, Orange, CA.
- Cowan, R., Cowan, W., & Swann, P. (2004). Waves in consumption with interdependence among consumers. *Canadian Journal of Economics*, 37(1), 149–177. doi:10.1111/j.0008-4085.2004.008_1.x
- Dohmen, T., Falk, A., Fliessbach, K., Sunde, U., & Weber, B. (2011). Relative versus absolute income, joy of winning, and gender: Brain imaging evidence". *Journal of Public Economics*, 95(3–4), 279–285. doi:10.1016/j.jpubeco.2010.11.025
- Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, 46, 687–724. doi:10.1016/S0014-2921(01)00208-2
- Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, 19(3), 291–313. doi:10.1287/isre.1080.0193
- Hennig-Thurau, T., Gwinner, K., Walsh, G., & Gremler, 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), 38–52. doi:10.1002/dir.10073
- Jeacle, I., & Carter, C. (2011). In TripAdvisor we trust: Rankings, calculative regimes and abstract systems. *Accounting, Organizations and Society*, 36(4–5), 293–309. doi:10.1016/j.aos.2011.04.002
- Jeong, E., & Jang, S. S. (2011). Restaurant experiences triggering positive electronic word-of-mouth (eWOM)

- motivations. *International Journal of Hospitality Management*, 30(2), 356–366. doi:10.1016/j.ijhm.2010.08.005
- Kemper, T. (1991). Predicting emotions from social relations. *Social Psychology Quarterly*, 54(4), 330–342.
- Kim, E. E. K., Mattila, A. S., & Baloglu, S. (2011). Effects of gender and expertise on consumers' motivation to read online hotel reviews. *Cornell Hospitality Quarterly*, 52(4), 399–406. doi:10.1177/1938965510394357
- Kollock, P. (1999). The production of trust in online markets. *Advances in Group Processes*, 16, 99–123.
- Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4), 673–707. doi:10.1162/jeea.2006.4.issue-4
- Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K., & Carnot, C. G. (1993). Attitude strength: One construct or many related constructs? *Journal of Personality and Social Psychology*, 65(6), 1132–1151. doi:10.1037/0022-3514.65.6.1132
- Lampel, J., & Bhalla, A. (2007). The role of status seeking in online communities: Giving the gift of experience. *Journal of Computer Mediated Communication*, 12(2), 434–455. doi:10.1111/jcmc.2007.12.issue-2
- Lee, H. A., Law, R., & Murphy, J. (2011). Helpful reviewers in TripAdvisor, an online travel community. *Journal of Travel & Tourism Marketing*, 28(7), 675–688. doi:10.1080/10548408.2011.611739
- Lee, J., Park, D.-H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, 7(3), 341–352. doi:10.1016/j.elerap.2007.05.004
- Li, Z., Huang, K. W., & Cavusoglu, H. (2012, December 15–16). *Can we gamify voluntary contributions to online Q&A communities? Quantifying the impact of badges on user engagement*. Proceedings of 2012 Workshop on Information Systems and Economics (WISE 2012), Orlando, FL.
- Liu, X., Schuckert, M., & Law, R. (2014). Can response management benefit hotels? Evidence from Hong Kong hotels. *Journal of Travel & Tourism Marketing*. doi:10.1080/10548408.2014.944253.
- Mauri, A. G., & Minazzi, R. (2013). Web reviews influence on expectations and purchasing intentions of hotel potential customers. *International Journal of Hospitality Management*, 34, 99–107. doi:10.1016/j.ijhm.2013.02.012
- Meng, X., Webster, S. A., & Butler, B. S. (2013, August 15–17). *Motivational effects of badge systems on participation in stack exchange social Q&A online community*. Proceedings of the Nineteenth Americas Conference on Information Systems, 2013, Chicago, IL.
- Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810), 56–63. doi:10.1126/science.159.3810.56
- Mkono, M. (2012). A netnographic examination of constructive authenticity in Victoria falls tourist (restaurant) experiences. *International Journal of Hospitality Management*, 31(2), 387–394. doi:10.1016/j.ijhm.2011.06.013
- Moldovanu, B., Sela, A., & Shi, X. (2007). Contests for status. *Journal of Political Economics*, 115(2), 338–363. doi:10.1086/518010
- 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.
- Öğüta, H., & Cezara, A. (2012). The factors affecting writing reviews in hotel websites. *Procedia-Social and Behavioral Sciences*, 58, 980–986. doi:10.1016/j.sbspro.2012.09.1077
- Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research*, 17(4), 392–414. doi:10.1287/isre.1060.0106
- Perretti, F., & Negro, G. (2006). Filling empty seats: How status and organizational hierarchies affect exploration versus exploitation in team design. *Academy of Management Journal*, 49(4), 759–777.
- Rustichini, A. (2008). Dominance and competition. *Journal of the European Economic Association*, 6(2–3), 647–656. doi:10.1162/JEEA.2008.6.2-3.647
- Schuckert, M., Liu, X., & Law, R. (2014). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*. doi:10.1080/10548408.2014.933154.
- Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing*, 21(4), 76–94. doi:10.1002/dir.20090
- Shah, C., Oh, J. S., & Oh, S. (2008). Exploring characteristics and effects of user participation in online social Q&A sites. *First Monday*, 13, 9. doi:10.5210/fm.v13i9.2182
- Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. C. (2002). Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, 136(2), 253–263. doi:10.1016/S0377-2217(01)00113-8
- Sparks, B. A., & Browning, V. (2010). Complaining in cyberspace: The motives and forms of hotel guests' complaints online. *Journal of Hospitality Marketing & Management*, 19(7), 797–818. doi:10.1080/19368623.2010.508010
- Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32(6), 1310–1323. doi:10.1016/j.tourman.2010.12.011
- Thompson, M. (2005). Structural and epistemic parameters in communities of practice. *Organization Science*, 16(2), 151–164. doi:10.1287/orsc.1050.0120

- Thye, S. R. (2000). A status value theory of power in exchange relations. *American Sociological Review*, 65(3), 407–432.
- Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, 30(1), 123–127. doi:10.1016/j.tourman.2008.04.008
- Weimann, G., Tustin, D. H., Vuuren, D. V., & Joubert, J. P. R. (2007). Looking for opinion leaders: Traditional vs. modern measures in traditional societies. *International Journal of Public Opinion Research*, 19, 173–190. doi:10.1093/ijpor/edm005
- Wetzer, I. M., Zeelenberg, M., & Pieters, R. (2007). Never eat in that restaurant, I did!: Exploring why people engage in negative word of mouth communication. *Psychology & Marketing*, 24(8), 661–680. doi:10.1002/(ISSN)1520-6793
- Willemsen, L. M., Neijens, P. C., & Bronner, F. (2012). The ironic effect of source identification on the perceived credibility of online product reviewers. *Journal of Computer-Mediated Communication*, 18(1), 16–31. doi:10.1111/j.1083-6101.2012.01598.x
- Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, 27(2), 634–639. doi:10.1016/j.chb.2010.04.014
- Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, 36(3), 6527–6535. doi:10.1016/j.eswa.2008.07.035
- Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology & Tourism*, 10(4), 283–295. doi:10.3727/109830508788403114
- Yoo, K. H., & Gretzel, U. (2010). Antecedents and impacts of trust in travel-related consumer-generated media. *Information Technology & Tourism*, 12(2), 139–152. doi:10.3727/109830510X12887971002701
- Yoo, K.-H., & Gretzel, U. (2011). Influence of personality on travel-related consumer-generated media creation. *Computers in Human Behavior*, 27(2), 609–621. doi:10.1016/j.chb.2010.05.002
- Zhang, R., & Tran, T. (2010). Helpful or unhelpful: A linear approach for ranking product. *Journal of Electronic Commerce Research*, 11(3), 220–230.
- Zhang, Z., Ye, Q., Law, R., & Li, Y. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, 29(4), 694–700. doi:10.1016/j.ijhm.2010.02.002
- Zhang, Z., Ye, Q., Zhang, Z., & Li, Y. (2011). Sentiment classification of Internet restaurant reviews written in Cantonese. *Expert Systems with Applications*, 38(6), 7674–7682. doi:10.1016/j.eswa.2010.12.147
- Zhang, Z., Zhang, Z., Wang, F., Law, R., & Li, D. (2013). Factors influencing the effectiveness of online group buying in the restaurant industry. *International Journal of Hospitality Management*, 35, 237–245. doi:10.1016/j.ijhm.2013.06.012

SUBMITTED: September 17, 2014

FINAL REVISION SUBMITTED:

March 3, 2015

ACCEPTED: June 15, 2015

REFEREED ANONYMOUSLY