"Status Effect" in User-Generated Content: Evidence from Online Service Reviews

Completed Research Paper

Christian Janze
Goethe University Frankfurt
Theodor-W.-Adorno-Platz 4
60323 Frankfurt, Germany
janze@wiwi.uni-frankfurt.de

Michael Siering
Goethe University Frankfurt
Theodor-W.-Adorno-Platz 4
60323 Frankfurt, Germany
siering@wiwi.uni-frankfurt.de

Abstract

This paper provides first empirical evidence on the impact of reviewer status on the objectivity of his contributions in online communities. While previous research indicates that user-generated online reviews guide consumer decision making, little is known about drivers of the actual review generation process. By drawing on Functional Role Theory, we derive four research hypotheses covering the general research question of factors influencing the objectivity of service reviews. Utilizing a data sample covering 413,077 reviews posted over 12 years on www.TripAdvisor.com, we evaluate our research model. Our findings indicate that with increased user status, review objectivity increases. Thus, we contribute to theory by generalizing the so-called "Popularity Effect" to a multi-dimensional "Status Effect", which is more widely applicable (e.g. settings without users-follow-users relationships). Furthermore, our results enable practitioners to find their most valuable content-producers.

Keywords: Online Community, User-Generated Content, Status, Service Review, Functional Role Theory, Electronic Markets

Introduction

User-generated online reviews are an important asset for online retailers as they attract customers and directly influence product and service sales figures (Chevalier and Mayzlin 2006; Forman et al. 2008; Sparks and Browning 2011). Consequently, the question of what makes reviews helpful has become central for information systems researchers in order to understand which factors lead to increased review diagnosticity (Mudambi and Schuff 2010). Previous work found that review aspects such as review depth or a review's readability influence the perceived review helpfulness (Ghose and Ipeirotis 2011; Korfiatis et al. 2012).

Although the importance of user-generated online reviews as well as the question of what makes them helpful is well-recognized in IS literature (Mudambi and Schuff 2010), the question of why a specific online review is written in a specific manner is not thoroughly addressed yet. Nevertheless, as the writing style has shown to influence review helpfulness (Siering and Muntermann 2013b) and objective writing is more "likely to carry a sense of expertise" (Goes et al. 2014), online retailers might incentivize those users with a helpful writing style to contribute further reviews. Therefore, it is important to understand whether the writing style of users remains constant over time or whether their behavior changes with increasing experience in generating online reviews.

Goes et al. (2014) provide first evidence on behavioral changes of users contributing content in a social commerce setting. In their study, they refer to the Hawthorne Effect (Adair 1984), which postulates that somebody's knowledge to be part of an experiment "modifies [the] behavior from what it would have been without the knowledge" (Adair 1984). In other words, the Hawthorne Effect encompasses a behavioral change of individuals, which occurs when they are aware that they are observed by third parties. Goes et al. (2014) show in the context of an online community that if users become more popular, they provide more objective online product reviews (measured by a decrease in emotionality). However, their study is based on a platform which offers users the possibility to follow other users. They utilize these connections between users to measure a users' popularity. In contrast, we argue that such a behavioral change is caused by an increased status on the platform – which is a multi-dimensional construct (reviewer specific, third party specific and review specific) and not fully captured by a single metric as the number of followers. In addition, not every platform offers the possibility to follow other users, which leads us to the following research questions: 1) Does the proposed "popularity effect" hold true (i.e. is measurable) for platforms where it is not possible to follow other users; 2) if the proposed popularity effect is measurable, is it caused by an increase in status due to increased experience on the platform? In order to investigate these research questions, we analyze which factors drive the emotionality expressed in online service reviews. By means of a sample of 413,077 reviews posted from May 05th, 2003 to April 8th, 2015 on www.TripAdvisor.com, we analyze whether the user status on a platform influences the objectivity of online service reviews.

Our results reveal that an increase of a users' activity on the platform is associated with behavioral changes. First, our findings suggest that with an increased reviewing experience as well as increased information disclosure, the contributed online service reviews become more objective. Second, more positive social feedback as well as an increased review depth also lead to less emotional and thus more objective online service reviews. Consequently, we show that an effect similar to the "popularity effect" as observed by Goes et al. (2014) exists for online service reviews on platforms where users are not able to follow each other: if contributors have more status on a platform, they produce more objective content. Therefore, we extend the previous understanding of Goes et al. (2014) by introducing the more generalized "Status Effect". Our results are primarily relevant for online retailers as they help to identify users providing the most objective online service reviews and thus generating value for their customers and, in consequence, increasing future turnover and profit generated on their platforms.

The remaining portion of this paper is structured as follows: Section 2 presents the research background of this study as well our research model built upon related theories in the fields of review diagnosticity and online communities. Section 3 presents the research methodology applied, dataset acquisition procedures and variable operationalization. Section 4 outlines the results of our empirical study. Finally, section 5 provides concluding remarks and possible directions for future research in this area.

Background and Research Model

User-Generated Online Reviews and Review Diagnosticity

The communication among consumers by means of online reviews is a central characteristic of social commerce platforms. These websites can be defined as places "where people can collaborate online, get advice from trusted individuals, find goods and services, and then purchase them" (Liang and Turban 2011). Previous research underlined the impact of online reviews on consumer's purchase decisions, whereas related studies found that online reviews can attract consumers and influence consumer behavior in case of physical products (Forman et al. 2008; Zhu and Zhang 2010) and services (Sparks and Browning 2011; Sparks et al. 2013; Ye et al. 2011).

A growing research stream investigates the question which factors influence the perceived helpfulness of user-generated online reviews. In this context, an initial study by Mudambi and Schuff (2010) found that especially the review's depth as well as its extremity have an influence on review diagnosticity. Further research papers investigate the influence of readability (Korfiatis et al. 2012), product-related aspects (Siering and Muntermann 2013b) as well as specific emotions (Yin et al. 2014; Wu et al. 2011) on the perceived helpfulness of online reviews.

A common characteristic of previous studies covering user-generated online reviews is that they investigate which aspects influence review helpfulness. It is assumed that for identifying helpful online reviews and users contributing these reviews, online retailers have to take into account the different aspects of a specific review (e.g. review depth or review extermity). As this is appropriate to identify helpful reviews at a specific point in time (e.g. when presenting the most helpful reviews related to a specific product), this procedure is not ideal for long-term analyses of user behavior on social commerce platforms. In this case, behavioral changes might occur due to an increase in status on the social commerce platform. Consequently, our study closes this research gap by investigating whether such behavioral changes exist, by examining which factors influence these behavioral changes and by discussing the implications for theory and practice.

Popularity Effect in User-Generated Content

Goes et al. (2014) provide first insights into the determinants of the objectivity of online reviews. Referring to the Hawthorne effect (Adair 1984), which describes behavioral changes of individual's depending on their perceived level of observation, Goes et al. (2014) primarily investigate whether the number of ties between users, which are created by following other users on a platform, influences their behavior. Their findings reveal that more popular users provide less emotional and thus more objective online reviews (Goes et al. 2014). They utilize Functional Role Theory (Biddle 1986) to argue that users with a high number of ties are perceived as experts in the community, want to be conform with their expert role and thus provide more objective information rather than emotional rhetoric (Goes et al. 2014).

Nevertheless, we argue that next to the number of ties, other aspects are more suitable to measure the expert status, especially in communities where it is not possible to follow other users at all. According to Signaling Theory (Spence 1973, 1974), users send signals to other users in order to reduce the uncertainty related to their status. Especially in the social network context, users have been shown to send signals to appear as a valuable part of a community (Lampe et al. 2007; Siering and Muntermann 2013a). Consequently, sending related signals might be the initial step within the online community to be perceived as an expert. As follows, the provision of these signals alone might lead to a behavioral change according to Functional Role Theory, as users want to conform to the expectations related to their desired expert role within the community. Consequently, we close this research gap by examining the influence of different factors signaled by users and third parties on the level of emotionality and ultimately objectivity expressed within their online reviews.

Research Model

Figure 1 presents our research model used to explain the determinants of the level of emotionality expressed in online service reviews. Based on the Hawthorne effect (Adair 1984), Functional Role Theory (Biddle 1986), Signaling Theory (Spence 1973, 1974) and the findings by Goes et al. (2014), we argue that

the status of a user in an online community is not thoroughly defined by the number of other users following him. In contrast, we propose that it is a mutli-dimensional construct, which is not fully captured by a single metric. Focusing on Signaling Theory, we argue that users send signals to other users in order to reduce uncertainty related to their position on the platform and thus to increase their status within the community. Consequently, we focus on different signals indicating the status of a specific user in an online community: review experience, information disclosure, social feedback and review depth. Therefore, we cover reviewer specific, third party specific and review specific signals in our research model. Based on the assumption that a higher status within the community changes user behavior, we hypothesize that an increased signaling of status within the community has an impact on user behavior. User behavior is represented by the level of emotionality of the reviews provided (which is, according to Goes et al. (2014), defined as the opposite of objectivity). Therefore, an increased status leads to a decreased level of review emotionality and thus to higher online review objectivity. In the following, we provide our research hypotheses as well as the rationale behind them.

Figure 1. Research Model to Explain the Emotionality of User Generated Reviews
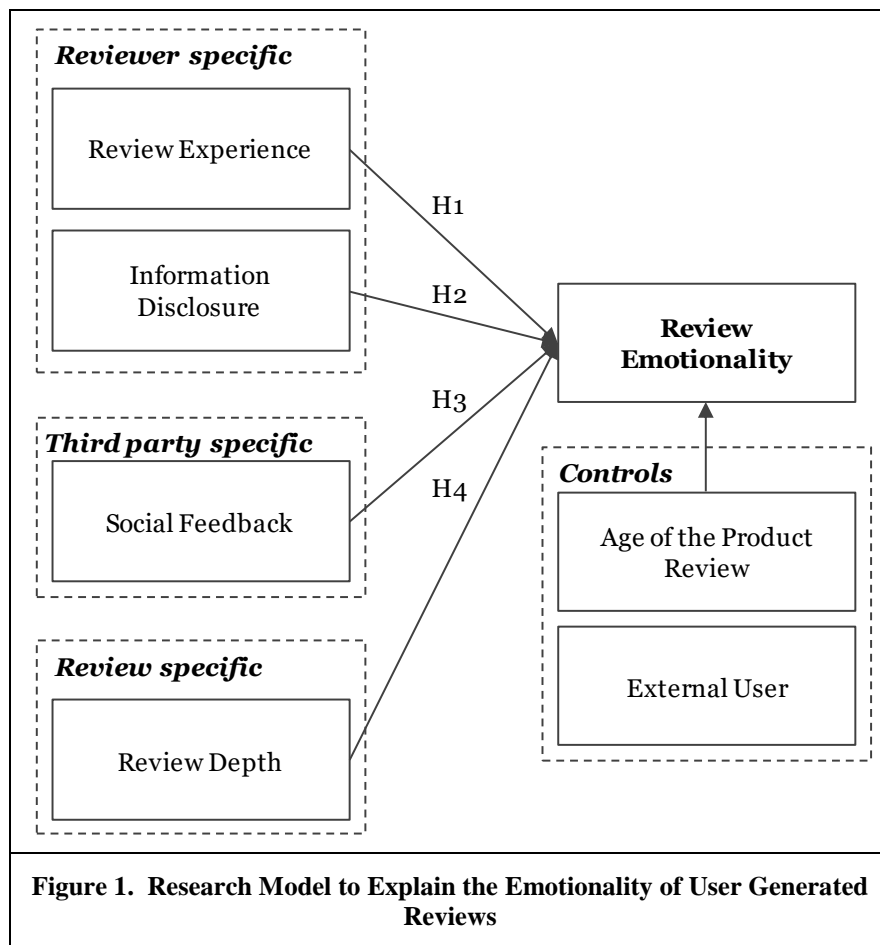
Review Experience

Review experience resembles the question of how experienced a specific reviewer is on a social commerce platform. Increased review experience can encompass the number of reviews provided on a platform or, more specifically, an increased contribution of online reviews related to a certain product or service category. We assume that an increased experience in writing online reviews is associated with learning effects regarding the ability to write helpful reviews and with a reduction in the number of reviews expressing spontaneous, undifferentiated perceptions about a specific service.

Furthermore, previous studies show that grading depends on the status of the person providing an evaluation (Moore and Trahan 1998). For instance, higher status of the person making the assessment leads to less positive evaluations (Kezim et al. 2005). As users can signal higher status within an online community with an increased number of contributions, it can be assumed that this also has an impact on the level of emotionality expressed in the online review. According to Functional Role Theory, a user acts according to his increased status and provides less emotional online reviews since this would be expected from an expert (Goes et al. 2014). As follows, we hypothesize:

H1: Increased review experience decreases review emotionality.

Information Disclosure

Information disclosure describes the level of information disclosed by a specific user within an online community (Norberg et al. 2007). Increased information disclosure can be seen as a signal to reduce the uncertainty perceived by other users – and thus as a factor positively influencing user status (Donath 2008; Lampe et al. 2007). With an increased amount of information available, other users can get a more precise impression about the real characteristics of a specific user. In this context, previous research has confirmed that there is a relation between the amount of information published and the real user personality (Marcus et al. 2006; Vazire and Gosling 2004).

As this higher level of non-anonymity might also lead to an increased status within the community, it can also be hypothesized that the Hawthorne effect in combination with Functional Role Theory applies to the level of information disclosure: We assume that users disclosing more information about themselves contribute more objective online reviews as they want to be perceived as experts in the community. We consequently hypothesize on the relation between information disclosure and review emotionality:

H2: Increased information disclosure decreases review emotionality.

Social Feedback

Many online review platforms allow users to evaluate the contributions of other users by voting on the perceived helpfulness of reviews (Mudambi and Schuff 2010). If a review platform displays the total number of helpfulness votes received by a specific user next to the online review, this represents an additional signal related to the status of the user contributing an online review. Furthermore, it was shown that social feedback influences user behavior, especially the level of content contribution (Siering and Muntermann 2013a; Hennig-Thurau et al. 2004).

Therefore, an increased level of social feedback within an online community displayed directly next to the online review is associated with an increased perception of responsibility for the online community. According to Functional Role Theory, it can be assumed that this increased level of responsibility also leads to the provision of more objective online reviews as the user behaves according to his role as an expert. Transferring this to the level of review emotionality, we hypothesize:

H3: Increased social feedback decreases review emotionality.

Review Depth

Review diagnosticity theory suggests that the provision of more comprehensive online reviews has a positive effect on the perceived helpfulness of a specific review (Mudambi and Schuff 2010; Korfiatis et al. 2012). An increased review depth consequently signals that a user has spent more time to provide a more thorough and precise evaluation. Thus, an impact of the general level of review depth on the level of review emotionality can be assumed: when a user wants to be perceived as an expert, he acts accordingly by providing more objective online reviews.

In addition, providing more comprehensive online reviews takes more time and forces the reviewer to take either more product or service characteristics into consideration and/or to think about them more extensively before posting the review. Consequently, this might also lead to a more objective writing style of the online review. As follows, we hypothesize on the relationship between review depth and review emotionality:

H4: Increased review depth decreases review emotionality.

Control Variables

Next to the different hypothesized relationships, we also include two important control variables into our research model. First, to control for time-dependent distortions of our results, we include the age of the review into our research model. This is especially important to control for possible status changes over time and for the fact that a specific user might provide different online reviews depending on the time he is active on the platform.

Second, platforms like www.TripAdvisor.com provide users with the possibility to contribute online reviews via third-party mobile apps such as Facebook. In this case, it is evident that users using third party apps do not have to visit the original website and can thus contribute online reviews more spontaneously. As higher spontaneity might lead to a less objective behavior, i.e. expressing more emotions due to emotional arousal, we include the usage of external apps as a control variable within our research model.

Research Methodology

Dataset Acquisition

The data used in this study was retrieved from TripAdvisor LLC (NASDAQ: TRIP). TripAdvisor is one of the largest providers of user-generated and travel-related reviews and opinions (comScore 2015). Founded in February 2000, TripAdvior is now operating in 45 countries, available in 28 languages and covering more than 200 million reviews, reaching an audience of 315 million unique visitors each month (TripAdvisor 2015a).

On April 8th, 2015, we acquired a large data set covering 413,077 user-generated restaurant reviews of 7,939 restaurants located in New York City (USA) with at least one review per restaurant written in English language. The acquired reviews were posted from May 05th, 2003 to April 8th, 2015.

Figure 2 presents an example of one of the restaurant reviews posted on TripAdvisor and used in our analysis. For each review, we extracted, depending on availability, a maximum of nine different data points (A-I). These can be divided into reviewer specific (A-E), third-party specific and review specific (G-I) ones:

Reviewer-specific data points are the reviewers' geographic location (A), the reviewers' title[1] assigned by TripAdvisor (B), the total number of reviews the reviewer has written (C), the total number of restaurant reviews (D) and the total number of cities the reviewer has written reviews in (E).

Furthermore, the total number of helpful votes the reviewer has received is provided as an evaluation of third-parties (F).

Review specific data points are the date of the review (G), the full-text of the review (H) and the number of helpful votes of the review (I). While review-specific data points are always available, reviewer-specific ones are not. For example, data point A is voluntarily provided by TripAdvisor users.

---

[1] Levels: No title, Reviewer, Senior Reviewer, Contributor, Senior contributor and Top contributor
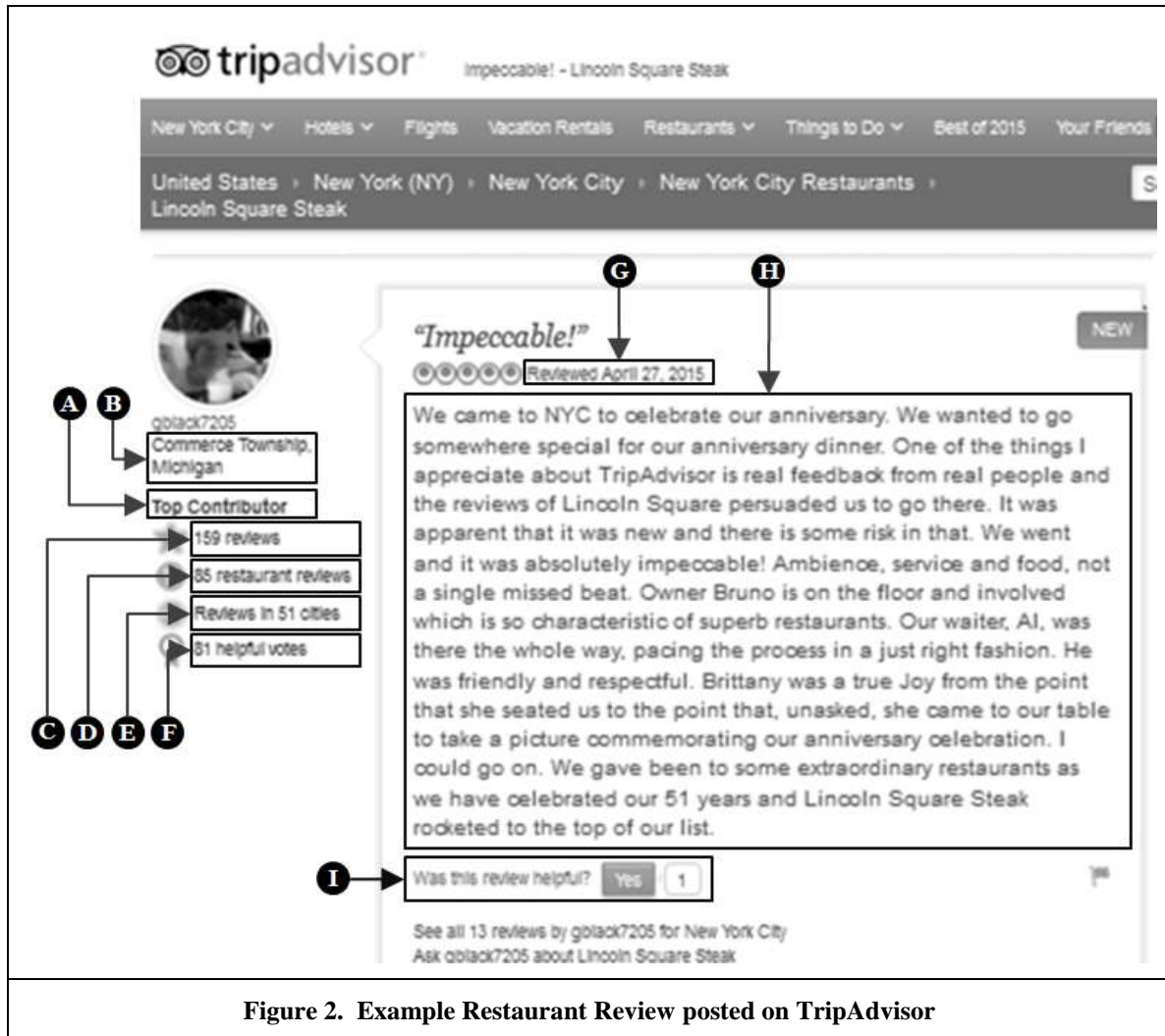
Figure 2. Example Restaurant Review posted on TripAdvisor

Variable Operationalization

In the following section, we present the variable operationalization of our research model introduced in Figure 1. While some of the variables presented in Table 1 (restRevs, citRevs, totRevs, hlpVotes, descWC) could be direcly used in our analysis, others had to be transformed and re-coded. In the following, starting with the emotionality variable, all transformations are described.

Review emotionality is calculated by means of automated content analysis on a dictionary-based approach (Krippendorff 2013). The wordlists used to determine the number of positive (desPos) and negative words (desNeg) in the description text contain 2,003 positive and 4,776 negative opinion words and have been used in numerous related studies (Hu and Liu 2004a; Liu et al. 2005; Hu and Liu 2004b). In addition, 23 negation words (e.g. no, not, can't, doesn't, etc.) are applied. However, it should be noted that a common problem of opinion and sentiment related analysis is that the presence of an positive (negative) opinion word does not automatically mean that the accompanying sentence is positive (negative) (Liu 2012).

Let desPosWC, desNegWC and desWC be the number of positive, negative and total words of a review's description text. Furthermore, let i denote a specific reviews index. Then, the emotionality of the description text *descEmot* can be written as:

$$desEmot_i = \frac{desPosWC_i + desNegWC_i}{desWC_i} \qquad (1)$$

To measure the emotionality of a user regarding a specific restaurant, we take into account his emotionality relative to the other users' evaluations of the same restaurant. This is necessary to account for the different qualities of the restaurants analyzed (e.g. fast food vs. fine dining). In order to rule out that the measured effects are only caused because of a correlation of positive emotions with positive restaurant ratings, we always consider a specific user's evaluation in comparison to other users' evaluations by calculating z-scores.

Each restaurant in our sample is indexed by a positive integer r, with $1 \leq r \leq 7{,}939$ and contains $1 \leq n$ reviews i. Thus, if μ_r denotes the mean and σ_r the corrected sample standard deviation of restaurant r's reviews emotionality, the absolute value of z-standardized emotionality scores can be written as:

$$emot_i = abs\left(\frac{descEmot_i - \mu_r}{\sigma_r}\right) \tag{2}$$

Thus, the transformed emotionality variable represents the unsigned number of standard deviations the observed desEmot variable deviates from the mean emotionality expressed towards a specific restaurant.

Table 1. Variables of a TripAdvisor Restaurant Review		
Variable		**Description**
Short Name	***Full Name***	
emot	Emotionality	Absolute value of the z-standardized ratio of positive and negative words in a reviews body divided by its word count
restRevs	Restaurant Reviews	Total number of restaurant reviews of reviewer
citRevs	City Reviews	Total number of cities a reviewer wrote reviews in
tripTit	TripAdvisor Title	Title that TripAdvisor's assigned to a reviewer
hasLoc	Has Location	Binary variable; containing 1 if the reviewer has voluntarily provided his geographic location and 0 otherwise
hlpVotes	Helpful Votes	Number of helpful votes of a specific review
corTotHlp	Corrected Total Helpful Votes	Corrected total helpful votes, defined as total number of helpful votes a reviewer received minus helpful votes of the review analyzed
totRevs	Total Number Reviews	Total number of reviews written by a reviewer
socFeed	Social Feedback	Ratio of corrected total helpful votes to the total number of reviews written by a reviewer
desWC	Description Word Count	Word count of the reviews' text
daysPassed	Days Passed	Age of the product reviews: Days passed until the data acquisition date
extUsr	External User	Binary variable containing 0 if the review was written on TripAdvisor and 1 if it was written on Facebook

The tripTit variable contains the re-coded TripAdvisor assigned reviewer title, which is primarily dependent on the reviewer's activity and defined as follows: 0=no title, 1=Reviewer, 2=Senior Reviewer, 3=Contributor, 4=Senior contributor, 5=Top contributor (TripAdvisor 2015b). The binary variable hasLoc yields 1 if a specific reviewer disclosed its geographic location and 0 otherwise. The corrected amount of helpful votes corTotHlp of a review is defined by subtracting the number of helpful votes a specific review received from the total number of helpful votes of the corresponding reviewer. By dividing the corrected number of helpful votes of a reviewer by the total numbers of reviews written by a reviewer, the social feedback variable socFeed is defined. The daysPassed variable contains the number of days passed since the publication of the review and the data acquisition date as of April 8th, 2015. As reviews on

TripAdvisor can be published via Facebook, the binary variable extUser contains 0 if the review was written on TripAdvisor and 1 if it was written on Facebook.

Regression Analysis

We use OLS regression for our analysis. Figure 3 presents an overview of the regression setup including the reviewer-, third party-, and review-specific dimensions of the status effect as well as the controls of our research model, as introduced in Figure 1.

Figure 3. Operationalization of Research Model

Note: Full variable names and descriptions can be found in Table 1.

While model A represents our base model, model B and C serve as robustness checks. Reviewer-specific dimensions are the review experience, operationalized in model A by the total number of restaurant reviews (restRevs) and information disclosure, measured by the presence of the voluntarily disclosed geographic location of the reviewer (hasLoc). The third-party specific dimension entails social feedback, which is measured by the ratio of the corrected number of helpful votes to the total number of reviews written by the reviewer (socFeed). The review specific dimension review depth is operationalized by the word count of the review description text (desWC). Controls are the membership duration, approximated by the number of days passed since the posting date of the review and the data acquisition on April 8th, 2015 (daysPassed) as well as the external user variable (extUsr), which differentiates between reviews written directly on TripAdvisor and indirectly by means of Facebook. The resulting regression equation of our Base Model A is presented in Equation 3, whereas Model B and C are formally given by Equation 4 and 5 respectively.

$$emot_A = \alpha + \beta_1 restRevs + \beta_2 hasLoc + \beta_3 socFeed + \beta_4 descWC + \beta_5 daysPassed \\ + \beta_6 extUsr + \varepsilon \tag{3}$$

$$emot_B = \alpha + \beta_1 citRevs + \beta_2 hasLoc + \beta_3 socFeed + \beta_4 descWC + \beta_5 daysPassed \\ + \beta_6 extUsr + \varepsilon \tag{4}$$

$$emot_C = \alpha + \beta_1 tripTit + \beta_2 hasLoc + \beta_3 socFeed + \beta_4 descWC + \beta_5 daysPassed \\ + \beta_6 extUsr + \varepsilon \tag{5}$$

Due to high correlations (see Table 3), only one variable expressing the experience a reviewer has acquired is included at the same time. Consequently, model A only contains the number of restaurant reviews, Model B the number of cities and Model C the title TripAdvisor assigned to the user.

Empirical Study

Descriptive Statistics

Table 2 presents summary statistics of the 413,077 reviews of restaurants located in New York City, which were posted on TripAdvisor from May 05th, 2003 to April 8th, 2015.

The absolute value of z-standardized emotionality scores (emot) yields a maximum (minimum) of 17.55 (0) standard deviations from the mean. The large difference of the mean (32.12) and median (15) number of restaurant reviews written by TripAdvisor users indicates the presence of some very active users. This observation is backed by the most active user, which alone contributed 1,564 restaurant reviews.

The same holds true for the mean and median number of cities in which TripAdvisor users wrote reviews in, which is 18.81 and 12 respectively with a maximum of 457. The binary variable hasLoc reveals that 87% of the reviewers disclosed their location in their TripAdvisor profile. Looking at the social feedback statistics, it is easy to see that the mean (median) number of helpful votes a reviewer received per review is 0.58 (0.46) with a maximum of 81. The mean number of words used in a review is 83.63, which is largely exceeded by the lengthiest review that contains 3,136 words. The oldest review was written 4,341 days before our data acquisition on April 8th, 2015. Looking at the binary variable extUsr, it is easy to see that only 4% of all reviews were written on Facebook.

Table 2. Summary Statistics of Restaurant Reviews (Cross-Sectional, by User and Review)						
Variable	**Minimum**	**Maximum**	**Mean**	**Median**	**Std. Dev.**	**N**
emot	0	17.55	0.70	0.55	0.70	411,440*
restRevs	0	1,564	32.12	15	53.86	413,077
citRevs	0	457	18.81	12	22.38	413,077
tripTit	0	5	3.41	4	1.64	413,077
hasLoc	0	1	0.87	1	0.34	413,077
socFeed	0	81	0.58	0.46	0.74	413,077
desWC	0	3,136	83.63	63.00	74.91	413,077
daysPassed	0	4,341	737.21	591.00	646.49	413,077
extUsr	0	1	0.04	0.00	0.20	413,077

Notes: As of April 8th, 2015. Full variable names and descriptions can be found in Table 1. *Due to cases where the corrected sample standard deviation with denominator 1-n yields zero for restaurants with only one review.

Pearson product-moment correlation coefficients of the nine variables used in our analysis are presented in Table 3. The total number of restaurant reviews written by a user (restRevs) is unsurprisingly highly correlated (0.786) with the number of cities a reviewer has written reviews in (citRevs). In addition the TripAdvisor assigned title of a user is highly correlated (0.602) with the number of cities a user has written reviews in. Consequently, these variables are only included separately within our models.

Table 3. Variable Correlations (Cross-Sectional)								
	emot	restRevs	citRevs	tripTit	hasLoc	socFeed	desWC	days Passed	extUsr
emot	1								
restRevs	-0.056	1							
citRevs	-0.079	0.786	1						
tripTit	-0.144	0.481	0.602	1					
hasLoc	-0.139	0.142	0.192	0.364	1				
socFeed	-0.058	0.003	0.079	0.120	0.138	1			
desWC	-0.165	0.086	0.110	0.122	0.146	0.128	1		
daysPassed	0.136	-0.077	-0.094	-0.267	-0.234	0.013	-0.058	1	
extUsr	0.234	-0.123	-0.173	-0.428	-0.526	-0.160	-0.178	0.569	1

Note: As of April 8th, 2015. Full variable names and descriptions can be found in Table 1.

Evaluation of the Research Model

To evaluate our research model, we performed three OLS regressions as shown in Table 4. The dependent variable is the absolute value of the z-score of review emotionality. While the first regression model (A) is our base setup, the second (B) and third model (C) represent robustness checks. To account for heteroscedasticity, White-corrected standard errors are used.

Table 4. Regression Analysis (n= 411,440 complete observations)					
* p<10%, ** p<5%, *** p<1% (White-corrected standard errors)					
	Model A (Base)		Model B (Robustness)		Model C (Robustness)	
	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value
(Constant)	0.7938	0.000***	0.8413	0.000***	0.7981	0.000***
restRevs	-0.0002	0.000***	-	-	-	-
tripTit	-	-	-0.0191	0.000***	-	-
citRevs	-	-	-	-	-0.0009	0.000***
hasLoc	-0.0235	0.000***	-0.0099	0.006**	-0.0200	0.000***
extUsr	0.6844	0.000***	0.6435	0.000***	0.6802	0.000***
socFeed	-0.0088	0.000***	-0.0064	0.000***	-0.0073	0.000***
desWC	-0.0012	0.000***	-0.0012	0.000***	-0.0012	0.000***
daysPassed	0.0000	0.000***	0.0000	0.000***	0.0000	0.000***
F-Value	5,235	0.000***	5,333	0.000***	5,266	0.000***
Adjusted R²	0.0709		0.0721		0.0713	

Note: Full variable names and descriptions can be found in Table 1.

First, research hypothesis H1 suggests that an increased review experience decreases review emotionality. Considering regression A as presented in Table 4, we can accept this hypothesis as the total number of restaurant reviews (restRev) has a negative effect on review emotionality, which is statistically significant at the 1% level. Furthermore, model B and C show that the same holds true if the review experience is

measured by the TripAdvisor assigned title (tripTit) and the number of cities a reviewer has written reviews in (citRevs).

Second, research hypothesis H2 states that increased information disclosure of a reviewer decreases review emotionality. Empirical results presented in Table 4 support this hypothesis, as the presence of the reviewer's location in the profile (hasLoc), is negatively associated with review emotionality. This effect is statistically significant at the 1% level. Again, robustness models B and C yield the same and statistically significant results at the 5% and 1% level respectively.

Third, research hypothesis H3 predicts that increased social feedback decreases review emotionality. According to our results presented in Table 4, we can accept this prediction for both the base model A as well as the robustness models B and C. The coefficient of the socFeed variable is always negative and statistically significant at the 1% level.

Fourth, research hypothesis H4 suggests that an increased review depth decreases review emotionality. Indeed, regression model A shows that the number of words used in a review decreases review emotionality, whereas this relationship is statistically significant at the 1% level. This observation holds true for robustness models B and C.

Taking into account the control variables, it can be easily observed that reviews posted via a third party app are more emotional (extUsr), which is statistically significant at the 1% level in Model A, B and C. This indicates that Facebook users tend to post more spontaneously. This results in more emotional and thus less objective reviews. Furthermore, it appears that the number of days passed since the review was written hast a slightly and statistically significant positive effect on review emotionality.

The adjusted R^2 of 0.0709 for regression model A shows that 7.09% of the variance are explained by our research model. Robustness models B and C yield very comparable results with an adjusted R^2 of 0.0721 and 0.0713. In addition, F-Values of 5,235 (Model A), 5,333 (Model B) and 5,266 (Model C) and their corresponding P-values suggest that the null hypothesis that every coefficient is zero can be rejected at the 1% level of significance.

As our sample covers a very large amount of online service reviews which might influence the statistical significance levels observed, we also apply a multitude of different analytical techniques described by Lin et al. (2013) to account for the p-Value problems arising from large sample sizes. First, we evaluate our model based on different sub-sample sizes. Therefore, we step-wise increase the size of random sub-samples drawn for the full sample from very small (1,000 observations) to very large (400,000 observations). Then, we calculate the p-values and the two-sided 95% confidence intervals for each variable in our model for each sample size. Furthermore, for comparing the results with the smaller sample size of Goes et al. (2014), we consider a sample size of 10% of our original full sample size. We draw 1,000 random sub-samples and investigate p-Value histograms for each variable as well as the 95% confidence intervals for each drawn sub-sample. The distribution of the p-Values as well as the bounds of the confidence intervals show that our results remain robust with much smaller sample sizes (and also different temporal distributions of online service reviews included).

Discussion

Our results indicate that reviewer specific, third party specific and review specific factors influence the writing style of online reviews. We show that the status of users within an online community influences the level of emotionality expressed in their reviews. Consequently, beyond an effect of general user popularity, we observe a more general "status effect" within our analysis that can be attributed to different dimensions. Building upon the Hawthorne effect and Functional Role Theory, users sending more signals in order to be perceived as an expert in the community also act according to their desired expert role by providing less emotional and thus more objective online service reviews.

Our results show that an increased number of restaurant reviews written by a reviewer, an increased number of cities a reviewer has written reviews in and a higher title TripAdvisor assigns to its users (used as proxies for the level of experience a user signals to increase his status) cause a decrease in review emotionality. Furthermore, we show that an increased level of information disclosure (representing a signal to reduce the uncertainty related to a specific user) is related to the provision of more objective reviews. In addition, we show that an increased level of social feedback of third parties and review depth

lower the level of emotionality and thus yield more objective reviews. This can also be explained by the fact that users perceive to have a higher status within the community and act accordingly by providing more objective reviews.

Taking into account the control variables leads to additional important insights: First, the usage of a third-party app is associated with a significant impact on the level of emotionality expressed in online service reviews. This positive influence of the usage of third-party apps on the level of emotionality might be explained by the fact that people using third-party apps to post online service reviews are more emotional. It can be assumed that they post the service review shortly after purchase, i.e. after visiting a restaurant and making a specific experience. Second, the age of the product review has a positive influence on the level of emotionality observed. This could be explained by the fact that internet users in general became more mature over the 12 years of our study and/or societal changes. Although a within-subjects design would be more appropriate for the observation of these behavioral changes, the sheer sample size of our study allows us to nevertheless draw conclusions about a large number of different users with different levels of experience and membership durations.

We are aware that our current study is confronted with several additional limitations. First of all, we focus on restaurant reviews to investigate service reviews. One could argue that the observed behavior might be different if other service categories are taken into account (e.g. less emotionally charged services). However, the current study does not analyze the general level of emotionality within a review. In contrast, we always take into account a user's deviation from the mean emotionality expressed towards a specific restaurant and thus account for the mean level of emotionality. As a consequence, we reduce the risk of an influence of the service type on the research results substantially.

Furthermore, because of the sheer amount of data available, we focus on restaurants that are geographically located in New York City. One could argue that the observed relationships are different for other regions in the world because of varying local preferences. However, we believe that New York City is a melting-pot of various cultures with a very active tourism industry and thus a representative proxy for users of www.TripAdvisor.com.

In addition, our study focuses on online service reviews. We are aware that (physical) products might be evaluated differently because of varying possibilities to assess them before and after purchase. Furthermore, when compared to product reviews (e.g. as in Mudambi and Schuff 2010), the service reviews investigated within our study are on average only half as long as the typical online product review. Consequently, we plan to evaluate the status effect using online product reviews in future research.

From a methodological point of view, we are aware that our current analysis does not directly take into account the specific quality of a certain restaurant by means of its star rating. One could argue that the restaurant's quality also has an influence on the emotionality expressed within the online review. We elaborated on including a reviews' star rating respectively its deviation from the average star rating of a restaurant in our model. However, from a theoretical perspective, it cannot be argued for sure whether there is an influence of the star rating on the level of emotionality or vice versa. Nevertheless, the inclusion of the z-standardized star ratings as an additional independent variable in our regression setup did not change our results. Furthermore, potential relationships between the quality of a restaurant and the level of emotionality are ruled out as we take into account the deviation of the average emotionality expressed towards restaurants. Consequently, potential biases are ruled out within our analysis.

Conclusion

User-generated online reviews are increasingly important for online retailers and online consumers. Whereas many previous studies investigated the helpfulness of online reviews, the question of what makes them objective and thus more valuable has been neglected in previous research. Based on a sample of restaurant reviews posted on www.TripAdvisor.com, we investigate whether aspects related to the review, the reviewer and to third parties influence the emotionality of reviews.

Our study contributes to the research stream of online reviews. In contrast to previous studies, which only investigate the question of what makes these reviews helpful or whether they have an impact on sales, we focus on a novel important aspect of online service reviews: their level of emotionality. Thereby, we extend the previous understanding that increased popularity can reduce the level of emotionality if users

are able to follow each other. In a context where the following feature is not available, we find that the level of emotionality depends on the status of a user in an online community. We observe that especially review experience, information disclosure, social feedback and review depth have a negative influence on review emotionality – and, in consequence, lead to higher review objectivity. Thereby, we also contribute to the literature on the Hawthorne effect and on Functional Role Theory by applying these theories in the context of online reviews and by explaining the observed behavior by these theoretical surroundings.

From a practical perspective, our findings are especially relevant for online retailers and their customers. For online retailers, our results allow for the identification of users which generate objective content. This information could be used to incentivize such users for the provision of further reviews. This is especially helpful for online retailers as online service reviews represent an important asset in relation to customer acquisition and retention. For customers, these results are relevant as well since customers are usually confronted with an enormous amount of online reviews resulting in an information overload. The insights from this study allow them to screen online service reviews more efficiently. In this context, customers can quickly assess the quality of a reviewer and decide of whether a specific service review is worth to be read or not. To do so, they should especially consider online service reviews of experienced users (e.g. many restaurant reviews written) with a high level of information disclosure (e.g. disclose their location), positive social feedback (e.g. many likes per review) and that write exhaustive online service reviews.

Within future research, we plan to extend our analyses in multiple directions. First of all, as another robustness check, we plan to include other cities as well as physical products instead of services. This would enable us to further investigate and confirm the observed relationships on the basis of an extended dataset. Such a dataset could also include additional socio-economical user characteristics such as gender. Besides, it is an interesting result that the question of whether a third party app is used for providing online service reviews has an influence on the level of emotionality expressed. Within future work, a more thorough analysis of the behavior in online communities using third party apps could consequently be performed. Finally, future research could also perform predictive analytics in order to automatically identify users providing objective reviews in order to foster their contribution behavior.

References

Adair, J. G. 1984. "The Hawthorne effect: A reconsideration of the methodological artifact.," *Journal of Applied Psychology* (69:2), pp. 334–345.

Biddle, B. J. 1986. "Recent Development in Role Theory," *Annual Review of Sociology* (12), pp. 67–92.

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.

comScore 2015. "comScore Ranks the Top 50 U.S. Digital Media Properties for March 2015," *http://www.comscore.com/Insights/Market-Rankings/comScore-Ranks-the-Top-50-US-Digital-Media-Properties-for-March-2015 (accessed: 2015/05/04)* .

Donath, J. 2008. "Signals in Social Supernets," *Journal of Computer-Mediated Communication* (13), pp. 231–251.

Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), pp. 291–313.

Ghose, A., and Ipeirotis, P. G. 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498–1512.

Goes, P. B., Lin, M., and Yeung, C.-m. A. 2014. ""Popularity Effect" in User-Generated Content: Evidence from Online Product Reviews," *Information Systems Research* (25:2), pp. 222–238.

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.

Hu, M., and Liu, B. 2004a. "Mining and Summarizing Customer Reviews," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Seattle, Washington, USA* .

Hu, M., and Liu, B. 2004b. "Mining opinion features in customer reviews," *Proceedings of the National Conference on Artificial Intelligence.*

Kezim, B., Pariseau, S. E., and Quinn, F. 2005. "Is Grade Inflation Related to Faculty Status?" *Journal of Education for Business* (80:6), pp. 358–364.

Korfiatis, N., García-Bariocanal, E., and Sánchez-Alonso, S. 2012. "Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content," *Electronic Commerce Research and Applications* (11:3), pp. 205–217.

Krippendorff, K. 2013. *Content Analysis: An Introduction to Its Methodology*, Los Angeles: Sage.

Lampe, C., Ellison, N., and Steinfield, C. 2007. "A Familiar Face(book): Profile Elements as Signals in an Online Social Network," *Proceedings of the Computer/Human Interaction Conference San Jose, California, USA*.

Liang, T.-P., and Turban, E. 2011. "Introduction to the Special Issue Social Commerce: A Research Framework for Social Commerce," *International Journal of Electronic Commerce* (16:2), pp. 5–13.

Lin, Mingfeng, Henry C. Lucas, and Galit Shmueli. 2013. "Research Commentary—Too Big to Fail: Large Samples and the P-Value Problem." Information Systems Research 24 (4): 906–917.

Liu, B. 2012. "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies* (5:1), pp. 1–167.

Liu, B., Hu, M., and Cheng, J. 2005. "Opinion Observer: Analyzing and Comparing Opinions on the Web," *Proceedings of the 14th International World Wide Web conference (WWW-2005), Chiba, Japan*.

Marcus, B., Machilek, F., and Schütz, A. 2006. "Personality in Cyberspace: Personal Web Sites as Media for Personality Expressions and Impressions," *Journal of Personality and Social Psychology* (90:6), pp. 1014–1031.

Moore, M., and Trahan, R. 1998. "Tenure Status and Grading Practices," *Sociological Perspectives* (41:4), pp. 775–781.

Mudambi, S. M., and Schuff, D. 2010. "What Makes a Helpful Online Review? A Study of Customer Reviews on amazon.com," *MIS Quarterly* (34:1), pp. 185–200.

Norberg, P. A., Horne, D. R., and Horne, D. A. 2007. "The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors," *Journal of Consumer Affairs* (41:1), pp. 100–126.

Sparks, B.A., Browning, V. (2011). "The impact of online reviews on hotel booking intentions and perception of trust," *Tourism Management* (32), pp. 1310–1323.

Sparks, B.A., Perkins, H.E., Buckley, R. (2013). "Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior." *Tourism Management* (39), pp. 1-9.

Siering, M., and Muntermann, J. 2013a. "How to Identify Tomorrow's Most Active Social Commerce Contributors? Inviting Starlets to the Reviewer Hall of Fame," *Proceedings of the 34th International Conference on Information Systems, Milan, Italy*.

Siering, M., and Muntermann, J. 2013b. "What Drives the Helpfulness of Online Product Reviews? From Stars to Facts and Emotions," *Proceedings of the 11th International Conference on Wirtschaftsinformatik, Leipzig, Germany*.

Spence, M. 1973. "Job Market Signaling," *The Quarterly Journal of Economics* (87:3), pp. 355–374.

Spence, M. 1974. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*, Cambridge, MA, USA: Harvard University Press.

TripAdvisor 2015a. "TripdAdvisor FactSheet," *http://www.tripadvisor.co.uk/PressCenter-c4-Fact_Sheet.html (accessed: 2015/05/04)* .

TripAdvisor 2015b. "What are review badges?" *https://www.tripadvisorsupport.com/hc/en-us/articles/200613997-What-are-review-badges- (accessed: 2015/05/04)* .

Vazire, S., and Gosling, S. D. 2004. "e-Perceptions: Personality Impressions Based on Personal Websites," *Journal of Personality and Social Psychology* (87:1), pp. 123–132.

Wu, P., van der Heijden, H., and Korfiatis, N. 2011. "The Influences of Negativity and Review Quality on the Helpfulness of Online Reviews," *ICIS 2011 Proceedings*.

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), pp. 634-639.

Yin, D., Bond, S. D., and Zhang, H. 2014. "Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews," *MIS Quarterly* (38:2), pp. 539–560.

Zhu, F., and Zhang, X. 2010. "Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics," *Journal of Marketing* (74:2), pp. 133–148.