How to Identify Tomorrow's Most Active Social Commerce Contributors? Inviting Starlets to the Reviewer Hall of Fame

Completed Research Paper

Michael Siering
Goethe University Frankfurt
Grüneburgplatz 1
60323 Frankfurt, Germany
siering@wiwi.uni-frankfurt.de

Jan Muntermann
University of Göttingen
Platz der Göttinger Sieben 5
37073 Göttingen, Germany
muntermann@wiwi.uni-goettingen.de

Abstract

Social commerce contributors share their experiences of products and services, which is appreciated by consumers and online retailers. Since such user generated content is especially valuable for online retailers, they incentivize the most active contributors to provide further product reviews. Our paper aims to explore the question of which user characteristics can be used to identify contributors of valuable contents. This is especially relevant for newly registered users who have not extensively contributed yet. Drawing upon the literature on social information processing, signaling and communication theory, we explore how individual user characteristics published in the personal user profiles are associated with the actual contribution activity. Therefore, we analyze more than 30,000 user profiles from amazon.com. We find that information disclosure, emotiveness and problem-orientation are related to the contribution activity. Consequently, our results advance the understanding of who are the most active contributors and provide new implications for theory and practice.

Keywords: Electronic commerce, user generated content, product reviews, content analysis

Introduction

In the social commerce era, user generated content in form of online product reviews has become an important factor within the purchase decision process (Chevalier and Mayzlin 2006). Product reviews published on websites like amazon.com are seen as an essential source of product-related information and thus influence the consumers' decision to buy a certain good (Forman et al. 2008; Zhu and Zhang 2010). Consequently, online product reviews have also become a central asset for online retailers: such forms of content can increase the number of online shop visitors and, as a consequence, the number of products sold (Mudambi and Schuff 2010).

Against this background, online retailers have started to foster the contribution of user generated content. For instance, users with a high number of reviews written are provided with incentives to contribute further reviews which then might attract additional online consumers. In this context, some retailers offer monetary incentives (Utz 2009) or provide products free of charge that then have to be reviewed (Amazon 2013a) in order to increase the contribution of user generated content. In case of well-established online-retailers, identifying users to be incentivized is easy: information about the past product reviews of a reviewer can be used in order to determine whether a reviewer contributes useful information. Nevertheless, such a procedure is not appropriate to rank newly registered users, since in this case, no past online product reviews are available at all. However, these users might also contribute valuable online product reviews if they were incentivized appropriately. Furthermore, such a procedure might neither be appropriate for online shops with a lower maturity, i.e. which have been launched recently so that ranking based on reviews is not possible at all, or for online shops which do not have a helpfulness ranking mechanism for online product reviews. Nevertheless, such online shops would also profit from identifying potential active contributors to offer them appropriate incentives and to foster their contribution behavior.

Existing research in the context of online product reviews has mainly focused on explaining their diagnostic value as well as their impact on product sales. It has been figured out that factors like review depth, review extremity and product type have an impact on online product review helpfulness (Mudambi and Schuff 2010). In addition, a relation to product sales has been observed (Forman et al. 2008; Zhu and Zhang 2010). In contrast, the question of which users most actively contribute to social commerce sites has rarely been addressed until now. Few studies have examined by means of surveys which factors motivate online consumers to write product reviews. It has been found that social feedback, economic incentives or the pleasure to help other consumers are reasons to publish online product reviews (Hennig-Thurau et al. 2004; Cheung and Lee 2012). However, these studies neglect user characteristics related to their communication behavior (Gibb 1961; Hancock et al. 2007) that can be derived from the individual user profiles as well as the general level of information disclosed in the profiles. As found in previous research, such information can be useful to draw conclusions about a users' real personality (Gosling et al. 2007; Marcus et al. 2006; Vazire and Gosling 2004). Consequently, these characteristics might also be taken into account as variables potentially explaining the contribution activity. From a methodological point of view, focusing on perceptual measures as in the case of previous surveys bears the disadvantage that only a relatively small sample of users can be analyzed. Thus, such an approach is not adequate to identify all users that might actively contribute user generated content.

Against this background, this paper deals with the research question of which user characteristics derived from user profiles explain the contribution of user generated content. In order to address this research gap and to overcome the shortcomings of the existent literature, we propose a research model to describe which factors influence the amount of user generated content contributed. Therefore, we build upon social information processing theory (Walther 1992), signaling theory (Donath 2008) and communication theory (Mortensen 2008) and hypothesize that the level of information disclosure, the level of emotiveness as well as the level of problem-orientation have a positive influence on a users' contribution to user generated content platforms. Furthermore, based on previous studies, several control variables are included in the research model as well.

Since previous research has confirmed that user characteristics extracted from user profiles are related to a user's real personality (Marcus et al. 2006; Gosling et al. 2007), we perform an empirical study based on user profile information of 30,707 users acquired from amazon.com to validate our research model empirically. Therefore, we operationalize the different variables of interest based on the information

extracted from the profiles. Relying on this study setup, we are able to confirm a positive impact of the level of information disclosure, the level of emotiveness as well as the level of problem-orientation on the amount of user generated content contributed. Thus, we extend the literature on user generated content in social commerce and illustrate theoretical and practical implications.

The remainder of this paper is organized as follows. In section 2, we outline the background of our study, present the research model and derive our research hypotheses. Section 3 encompasses the research methodology applied, including dataset acquisition, variable operationalization and regression analysis. Our empirical study is presented and discussed in section 4. Finally, section 5 concludes.

Background and Research Model

User Generated Content in Social Commerce and Review Diagnosticity

In recent years, the generation and use of user generated content (UGC) has significantly increased, which goes back to "people who voluntarily contribute data, information, or media that then appears before others in a useful or entertaining way" (Krumm et al. 2008). This content has also found its way into classic electronic commerce web sites. Such social commerce sites differ from classic online shopping sites because consumers are enabled to communicate with each other and thereby influence the commerce process (Curty and Zhang 2011). Social commerce web sites can consequently 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).

Sharing product and service-related information and seeking advice from others before buying can be beneficial to consumers. This additional information can be relevant for the purchase decision, e.g. because of reduced search costs or better justified purchase decisions (Chevalier and Mayzlin 2006). Furthermore, retailers or service providers offering social commerce functionalities on their web sites can also benefit from customers who are willing to contribute UGC. For example, informative product reviews and recommendations can attract new consumers. As follows, such social commerce activities can contribute to increased profits (Curty and Zhang 2011). Consequently, major retailers such as amazon.com have offered their users the opportunity to contribute and to evaluate online product reviews.

As a result, analyzing the impact of such social commerce-related UGC, especially in the form of online product reviews, has become a vital research field. For example, existing research has addressed the question of whether such user generated product reviews do have a positive impact on product sales (Forman et al. 2008; Zhu and Zhang 2010). Other studies have focused on the question of how such product reviews are appreciated by others. Based on review diagnosticity theory, the related research explores the relevant factors that contribute to a review's helpfulness, i.e. what makes the difference between an unhelpful and a helpful review. In a study conducted by Mudambi and Schuff (2010), review extremity, review depth and product type have been found to have a significant impact on review helpfulness. Building upon these findings, subsequent studies have further explored the helpfulness construct and provided evidence that review readability (Ghose and Ipeirotis 2011; Korfiatis et al. 2012), specific review foci such as a focus on product qualities (Siering and Muntermann 2013) or expressed emotions (Yin et al. 2011; Wu et al. 2011) can affect perceived review helpfulness.

Contribution to User Generated Content Platforms in Social Commerce

Several previous studies have investigated which factors motivate users to participate and to contribute to user generated content platforms in the social commerce context. In a study focusing on users of different web-based opinion platforms, Hennig-Thurau et al. (2004) find that mostly desire for social interactions, desire for economic incentives, concerns for other consumers and the potential to enhance ones self-worth are motivations to participate in electronic word of mouth platforms. In a different survey-based study, Cheung and Lee (2012) confirm the importance of reputation, sense of belonging to a group as well as enjoyment to help other consumers as key motivational aspects for contributions. In contrast, Utz (2009) finds that, although increasing reputation is important for users contributing user generated content, altruism and the pleasure for interactions have increased importance. In a study comparing users of videoblogging and weblogging platforms, Stoeckl et al. (2007) highlight that for the contributors,

especially the production of videos is associated with fun. Finally, on a product level and by conducting a meta-analytic review of offline word-of-mouth communications, Matos and Rossi (2008) find that product-related aspects like satisfaction with a product or its perceived value are also important reasons for users to write a specific product evaluation.

To summarize, previous studies find that on the one hand, non-monetary incentives play a major role for users to contribute user generated content like product reviews. In this context, altruism, i.e. helping users without expecting a return, is an important motivation for contributions. On the other hand, the studies also highlight that social aspects (i.e. feedback from the group, social interactions) as well as economic incentives may motivate users to contribute.

A related stream of research deals with the analysis of factors that influence the contribution of knowledge within electronic networks. As illustrated by Ma and Agarwal (2007), self-presentation, i.e. the means by which a person presents herself online, influences if a person's identity is perceived to be verified. This in turn has a positive effect on knowledge contribution. Similar to the above stream of literature, Wasko and Faraj (2005) provide further evidence that individuals who perceive that their participation will positively affect their reputation, contribute to a network more actively.

In these different studies, the motivation to contribute to user generated content platforms is measured by means of survey instruments focusing on general user motivations. However, for online retailers and providers of user generated content platforms, the question of which user characteristics explain the amount of content published is more important since these characteristics given in the individual user profiles could be used to identify potential future contributors. These could then be given incentives (as identified by the previous studies) to post an increased amount of content. Consequently, we close this research gap by identifying user characteristics that indicate an increased participation in user generated content platforms. Users with a promising profile could then be identified by searching for these characteristics in the user profiles.

### *Extracting User Characteristics from User Profiles*

Extracting user characteristics from user profiles is based on the assumption that the characteristics reported reflect reality and that users do not manipulate the information given. In this context, different studies have evaluated whether user information published correlates with their real personality and whether such information can consequently be used to draw further conclusions on a users' personality.

In this context, a study by Gosling et al. (2007) takes into account facebook profiles and evaluates whether impressions evoked by the profile information given converge with how the users are seen by close acquaintances. Within this study, strong correlations between both variables are observed. These results are further supported by studies analyzing personal websites. Here, it is confirmed that the information published is valuable to draw valid conclusions about an authors' personality (Marcus et al. 2006; Vazire and Gosling 2004). Besides, Back et al. (2008) take into account different e-mail addresses and analyze whether these allow to draw conclusions about the personality of the corresponding owner. Here, it is found that to some degree, even this small piece of information is related to a users' real characteristics. Moreover, Lampe et al. (2007) confirm that the information disclosed within profile fields of facebook users is related to their number of friends. In addition, analyzing online dating platforms, Fiore et al. (2008) find that taking into account and analyzing free texts inserted by users is helpful to draw conclusions about a users' attractiveness.

Thus, focusing on user generated content platforms in the social commerce context, it can also be assumed that profile data, including free texts, is valuable to draw conclusions about user characteristics and consequently, user behavior in terms of contribution to social commerce sites. Analyzing such existing profile information for extracting user characteristics instead of using perceptual measures collected by means of surveys is linked with several advantages (Boyd et al. 1993). The data is in principle accessible to all researchers and thus, studies can be replicated more easily. Furthermore, there is no bias related to certain groups of users that do not take part in a survey – but which are nevertheless analyzed if profile information published online is taken into account (Boyd et al. 1993).

Research Model

In the following, we present our research model to explain which user characteristics determine the level of user generated content contributed in the social commerce context. Therefore, we build upon previous research focusing on the motivational aspects that lead to UGC contribution. Furthermore, our research model integrates social information processing theory (Walther 1992), signaling theory (Donath 2008) and communication theory (Mortensen 2008) to enhance the previous understanding. In our research model, we hypothesize that the level of information disclosed by reviewers, their level of emotiveness, as well as their problem-orientation have a positive influence on their level of UGC contribution (H1-H3). Since social feedback, as well as economic incentives, also influence the level of UGC contribution, we include these as control variables within the research model. Finally, since previous research has shown that gender has an influence on the perceptions of consumers within the electronic commerce context (van Slyke et al. 2002; Coley and Burgess 2003), we also include gender as control variable in the research model.

Figure 1. Research Model to Explain the Contribution of User Generated Content (UGC) in Social Commerce
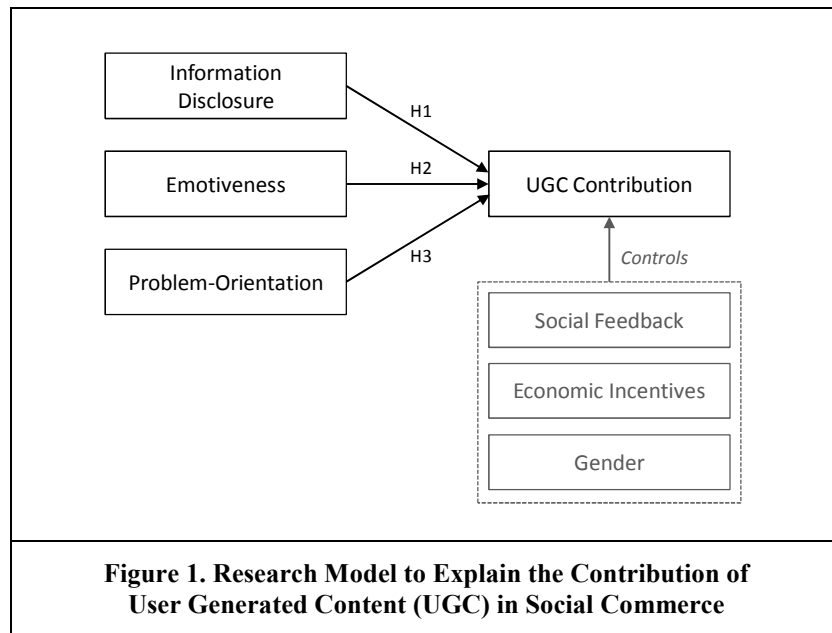
Level of Information Disclosure

In the context of social commerce, information disclosure can be defined as an activity of users who "actually provide personal details" (Norberg et al. 2007). Signaling theory (Spence 1974; Donath 2008) is closely related to the question of what information is provided within user profiles and why this information is posted. Here, signals are defined as "activities or attributes of individuals in a market which, by design or accident, alter beliefs of or convey information to, other individuals in the market" (Spence 1974). In a social commerce context, it is assumed that profile items represent such signals that "indicate hidden qualities of a person" (Donath 2008; Lampe et al. 2007) that otherwise cannot be observed directly and can thus increase trust in certain users. Some signals are assumed to be more reliable than others (Donath 2008; Lampe et al. 2007). Assessment signals are signals that are seen to be reliable since sending the signal requires that the corresponding quality is possessed (for instance, lifting a certain weight). In contrast, conventional signals are not seen as fundamentally reliable since "the link between signal and quality is arbitrary" (Donath 2008). One example is the actual age given in an online profile, where it is easy to fill in a different birth date if such a field is not validated by checking an identity card. Consequently, the more signals given within a user profile (and especially assessment signals like a name or e-mail address verified by the site operator as in the case of Amazon, where a user name is labeled as "real name" if it has been checked), the more reliable the profile can be seen (Lampe et al. 2007). Thus, users sending more signals in form of an increased amount of fields filled in their user profile try to be seen as a valuable part of the community. Next to sending signals by means of user

profiles, a complementary way for fostering such an increase in reputation within the social commerce context can consequently be seen in increasing the content contribution in general. With this respect, users might provide information about the product but also might provide further information about the circumstances of product usage. This might also serve as a signal and increase trust in the corresponding user. As a consequence, this also leads to the assumption that users sending more signals by means of information disclosed in their user profile also contribute more user generated content in form of online product reviews to the platform.

In addition, one primary motivation for contributing to user generated content platforms is pleasure for social interactions (Hennig-Thurau et al. 2004) as well as a desire for contacts (Stoeckl et al. 2007). In the social commerce context, such interactions occur to a lesser extent through discussions on the websites since the related websites list different online product reviews posted by consumers. Instead, social interactions may arise through direct communication between the users. In this context, the information signaled on the user profiles may be used by users to get in contact with other consumers. Most important, information on how to contact a certain user is central to be able to start a discussion. Exemplary profile information for that purpose could be a user's e-mail address or website, which is given in the user profile. Furthermore, information on the interests and reference points of a consumer can build a common ground with other consumers which can then foster interactions (Lampe et al. 2007).

As follows, in the social commerce context, users can mainly get the attention of other consumers as well as increase their reputation by writing online product reviews. In this case, other consumers can read the product review, comment on the product review as well as take a look at the corresponding user profile and the expressed signals in order to contact a user, for instance to ask clarifying questions. Consequently, we hypothesize a relation between the amount of personal information disclosed in user profiles and the amount of content contributed:

H1:   *The level of personal information disclosure has a positive impact on the contribution to user generated content platforms.*

Level of Emotiveness

Emotions can influence the behavior of individuals significantly and "motivate adaptive thought and action" (Izard 2002). Emotions also play an important role within the purchase decision process as well within the context of review diagnosticity theory. Thereby, previous research has shown that emotions are important when purchase decisions are made. For instance, in advertising research, diverse studies have found that emotional advertisements are published to evoke emotions at the customer level which might finally lead to purchase decisions (Sonnier et al. 2011). Moreover, in the context of review diagnosticity theory, it has also been shown that emotions influence the helpfulness of online product reviews (Siering and Muntermann 2013).

Furthermore, social information processing theory (Walther 1992) explains how emotions are expressed within computer mediated communications. In this case, it is assumed that emotions which are typically expressed non-verbally within offline communications are transmitted via verbal communication cues in online contexts (Walther 1992; Walther et al. 2005). Thus, it has been found that the language of people expressing emotions differs, for instance regarding punctuation used (Gill et al. 2008) as well as the number of words or messages published (Hancock et al. 2007). Consequently, it can be assumed that users being generally more emotional than other users also take the opportunity to express these emotions in their user profiles since user profiles have been found to represent real user characteristics (Gosling et al. 2007). Furthermore, it can also be assumed that these users express an increased amount of emotions via online product reviews, which might be reflected in the single reviews and, more important, in the amount of online product reviews contributed to user generated content platforms. As follows, we hypothesize:

H2:   *A user's expressed level of emotiveness has a positive impact on the contribution to user generated content platforms.*

Level of Problem-Orientation

In communication theory, problem-orientation represents a category of supportive communication expressing "a desire to collaborate in defining a mutual problem and in seeking its solution" (Gibb 1961), which positively affects communication behavior. We build upon this aspect of communication theory and hypothesize on the relationship between the level of problem-orientation expressed in a user's profile and her or his behavior to actively contribute to user generated content platforms.

Online product reviews play an important part in online consumers' purchase decisions since they can acquire product related information and evaluations independently from the product manufacturer before purchase (Chevalier and Mayzlin 2006). Thus, online product reviews help to support the online consumers' decision-making processes and help in solving the consumers' decision problem on whether to buy a certain product by giving information beyond advertisements (Tong et al. 2007). Thereby, the review helps during the different phases of the decision making process (Sprague 1980), i.e. from identifying the problem to implementing a solution, i.e. deciding on which product to buy.

In this context, helping other consumers with solving this decision problem is seen as a motivation for publishing user generated content on social commerce platforms (Hennig-Thurau et al. 2004; Cheung and Lee 2012). Such a motivation is closely related to the concept of altruism, i.e. helping other consumers without expecting a return (Smith 1981). Here, reviewers make available both positive and negative experiences with a product to provide decision support (Hennig-Thurau et al. 2004). In order to supply product reviews valuable to online consumers and to provide support for their purchase decisions, reviewers have to focus on typical product-related aspects satisfying the online consumers' information needs (Tong et al. 2007). Consequently, these product reviews can be assumed to be problem-specific. Thus, it can be expected that reviewers being more oriented on problem-solving in general as expressed in their user profile also provide more online product reviews in order to support other consumers' purchase decisions. As a consequence, we hypothesize:

H3:   A user's expressed level of problem-orientation has a positive impact on the contribution to user generated content platforms.

Control Variables

In user generated content platforms, users can often get approval by other consumers in form of helpfulness rankings. Such feedback represents a positive recognition from other users and could also symbolize to a user that she or he is perceived by other consumers as an expert (Hennig-Thurau et al. 2004). Such rankings can foster a positive reputation which has been shown to be a main motivational factor to contribute to related online communities (Kollock 1999). In addition, such feedback by other consumers is also perceived as "ego strokes" (Utz 2009) and can thus increase the amount of user generated content contributed. As a consequence, we include the amount of social feedback as control variable within our research model:

C1:   Social feedback has a positive impact on the contribution to user generated content platforms.

Providers of electronic word of mouth platforms can reward users publishing content on the platform by means of monetary incentives. For instance, some communities pay their users rewards if their postings are perceived to be helpful (Utz 2009) or provide users with products free of charge for reviewing them (Amazon 2013a). Economic incentives are seen as a form of approval for users contributing to such platforms (Hennig-Thurau et al. 2004) and economic incentives have been recognized as an important determinant of human behavior (Lawler 1984). Consequently, these can also increase the quality of product reviews written (Wang et al. 2012), so we add the provisioning of economic incentives as an important motivation for contributions to user generated content platforms as control variable within our research model:

C2:   Economic incentives have a positive impact on the contribution to user generated content platforms.

In the context of electronic commerce, previous research has also focused on the question of whether there are differences regarding the shopping behavior of men and women. In this context, it has been found that there are differences in product purchase frequency (Coley and Burgess 2003) that women are

less technical and spontaneous during the purchase process (Dholakia and Chiang 2003) and that women are more likely to favor a physical shopping experience (Hui and Wan 2007). Furthermore, men have been shown to purchase more likely from online shopping sites (van Slyke et al. 2002). Consequently, since men are more likely to buy online, one might also assume that men are more likely to contribute to user generated content platforms. Although these differences may have diminished over time (Hernández et al. 2011), we include gender as a control variable within our research model:

C3:   *Men contribute more to user generated content platforms than women.*

Research Methodology

Dataset Acquisition

To test our research model empirically, we acquire a dataset composed of user profile information from amazon.com. Based on Mudambi and Schuff (2010), we first select six different product categories and browse the top 100 best selling products for each category. The product categories selected cover products that can be easily evaluated before purchase (search goods) and products that have to be tried out before an evaluation is possible (experience goods). We include the categories *Camera & Photo*, *Computer Printers* as well as *Cordless Telephones* representing search goods and *MP3 Players*, *Music* as well as *PC-compatible Games* representing experience goods. For each product, we then browse the product reviews published and crawl the corresponding user profiles.

Many users contribute more than one online product review, so the users covered have also published product reviews related to other products beyond the six pre-defined categories. Consequently, we ensure that our sample of user profiles covers users publishing reviews related to diverse search and experience goods. This is important since reviews related to both product categories differ, which might also apply to the different users and could consequently bias the results if a certain product category would be neglected. Additionally, this procedure ensures that our sample covers users with different quantities of product reviews published which would not be the case if we only crawled the Amazon top reviewers list instead.

To be able to conduct our analyses, we take into account profiles of users whose reviews received at least one helpfulness vote. This is necessary to be able to calculate the percentage of helpful votes on a user basis. Furthermore, to be able to determine the gender of the different users, we only take into account those profiles containing a user name that matches a comprehensive pre-defined word list consisting of different forenames from the U.S. Social Security Administration (Social Security Administration 2011). Nevertheless, the results of our study remain robust if both selections are not performed. However, no gender impact as well as no impact of social feedback can be measured in this case.

Variable Operationalization

The operationalization of the variables is conducted on the basis of the user profiles collected. This is done directly by calculating measures of whether certain information is provided or not, or indirectly by extracting the required information from textual profile components by means of content analysis. Table 1 provides an overview of the different variables.

Dependent Variables

As dependent variables measuring the level of contribution of user generated content, we extract the total number of reviews a user has posted (*no_reviews*). In addition, we also extract the Amazon user rank. This variable takes into account the amount and the helpfulness of the different reviews posted, weighted by the time that has passed since a review has been posted. Thus, recent online product reviews become more important (Amazon 2013b). To reduce the differences between the particular ranks, we take into account the logarithm of the rank (*ln_rank*). Consequently, we ensure that we measure both the amount of reviews as well as the amount of helpful reviews posted.

Main Independent Variables

Related to the independent variables, we measure three indices representing the level of information disclosure as proposed by Lampe et al. (2007). Therefore, we first extract the following dummy variables from the user profiles that are set to one if the corresponding information is given: e_mail, website, birthday, real_name, interests, shared_recent_purchases, photo, wishlist, slogan, in_my_own_words as well as location. Then, we calculate three information disclosure indices according to Lampe et al. (2007): the *contact_index* encompasses profile fields indicating the "willingness to share off-site connections with others", the *interest_index* expressing "personal preferences and self-descriptive information" as well as the *referent_index* covering "common points of reference among users". Each index is determined by calculating the average of the variables included as described in Table 1.

While these three indices representing the level of information disclosure can be directly calculated based on the user profile data (given the (non-)existence of the above mentioned data fields), the concepts of emotiveness and problem-orientation require a more in-depth analysis of the profile data. We therefore decide to analyze the free text that can be added to the profile by the individual user. The free text fields taken into account comprise "in my own words", "interests" and "slogan", which can be edited by the user. For example, the currently highest ranked reviewer (#1 Hall of Fame Reviewer[1]), explains "in my own words" that she is "An engineer by day and a photographer in my spare time. I always enjoyed sharing my experience with various products and found that attention to details that is required in my day job helps me write product reviews. I depend on Amazon reviews for my own purchases and I hope people will find my contributions helpful in their decisions".

Thus, we have analyzed these textual data fields by means of content analysis methodology which can be defined as "systematic, objective, quantitative analysis of message content" (Neuendorf 2002). Content analysis is usually applied to textual data in order to draw conclusions about the author or about communication content and style (Rosenberg et al. 1990). Generally, content analysis can be performed in two alternative ways (Rosenberg et al. 1990): First, one can analyze the textual data on the basis of human coding, which requires human coders who manually read and code the content. Second, one can follow an automated approach, i.e. computer coding, which requires pre-defined dictionaries that correspond to the variables of interest and a method of applying them (Neuendorf 2002).

We decide to follow a computer-based approach based on pre-defined dictionaries to operationalize emotiveness and problem-orientation. In contrast to manual coding of textual data, this approach has no problem with inter-coder reliability. In addition, since well-respected dictionaries are publicly available, results can be reproduced easily without any loss in quality (Weber 1984). Further, computer coding enables us to analyze massive data volumes (in our case user profiles). The computer-based approach thereby enables us to perform the data coding within an acceptable period of time.

During computer coding, the single words of a text are mapped to pre-defined classes representing psychological categories. For that purpose, dictionaries related to these psychological categories are applied to calculate frequencies of how often the related words contained in the dictionary can be found within the analyzed text data. In our study, we make use of the Harvard IV-4 dictionaries provided by the General Inquirer (GI) text analysis framework (Stone et al. 1962). These dictionaries are well-established in the scientific literature, are being applied since decades until today and provide standardized classifications due to extensive previous application and validation (Weber 1990).

In order to assess emotiveness and problem-orientation, we apply the corresponding word lists from the Harvard IV-4 dictionary. First, we make use of the *"positiv"* and *"negativ"* word lists. These word lists contain words related to positive and negative emotions and are used to assess the emotiveness expressed within the profile fields analyzed. The same is done for problem-orientation by applying the *"solve"* word list, which refers to mental processes associated with problem solving. For emotiveness and problem-orientation, we then calculate a ratio of the number of words related to the word list divided by the total number of words. This leads to *emotion* and *solve* variables for each different profile field showing values

---

[1] http://www.amazon.com/review/top-reviewers (accessed: 05/03/2013);
http://www.amazon.com/gp/pdp/profile/A2D1LPEUCTNT8X/ (accessed: 05/03/2013)

between 0 and 1. In a final step, we then take the average of each variable related to the three different profile fields analyzed and multiply the result by 100.

Control Variables

In addition, we operationalize our control variables. The percentage of helpfulness votes a user received for his reviews represents the social feedback obtained from other users. Therefore, we both extract the number of helpfulness votes as well as the number of total votes received. The percentage of helpfulness votes is then determined as the fraction of the helpfulness votes divided by the total votes (*helpful_percent*). Furthermore, the dummy variable *vine_voice* represents economic incentives received: the variable is set to one if the profile indicates that a user takes part in the Amazon vine program that offers free products to the reviewer. Finally, we extract the user name displayed to be able to determine the gender of the corresponding user. Therefore, we make use of lists containing male and female forenames acquired from the U.S. Social Security Administration (Social Security Administration 2011). The list contains the forenames as well as their occurrences for men and women. The gender of each user is then determined by analyzing for which gender a certain forename is given most frequently.

Table 1. Operationalization of Independent (IV) and Dependent Variables (DV)			
Variable Type	Research Hypothesis	Variable	Operationalization
IV	H1: Information Disclosure	contact_index	Average of the dummy variables covering the profile fields e_mail, website, birthday and real_name.
		interest_index	Average of the dummy variables covering the profile fields interests, shared_recent_purchases, photo, wishlist, slogan and in_my_own_words.
		referent_index	Equal to the dummy variable location.
	H2: Emotiveness	emotiveness	Average of the emotiveness of the profile fields "in my own words", "interests" and "slogan". Each score is calculated as the number of positive and negative words divided by the number of total words multiplied by 100.
	H3: Problem-Orientation	solve	Average of problem-solving related terms expressed in the profile fields "in my own words", "interests" and "slogan". Each score is calculated as the number of words related to problem solving by the number of total words multiplied by 100.
	C1: Social Feedback	helpful_percent	Number of helpful votes per user divided by the total number of votes received.
	C2: Economic Incentives	vine_voice	Indication whether the user participates in the Amazon vine program. 1 if yes, 0 if no.
	C3: Gender	male	Gender extracted from the user name. 1 if male, 0 if female.
DV		no_reviews	Number of reviews posted.
		ln_rank	Logarithm of the Amazon user rank.

Regression Analysis

In order to test the research hypotheses, we run two different ordinary least squares (OLS) regressions, both taking into account the independent variables as defined in table 1, whereas *controls* represents the different control variables.

$$no\_reviews = \alpha + \beta_1 contact\_index + \beta_2 interest\_index + \beta_3 referent\_index + \beta_4 emotiveness + \beta_5 solve + \beta_6 controls + \varepsilon \tag{1}$$

$$ln\_rank = \alpha + \beta_1 contact\_index + \beta_2 interest\_index + \beta_3 referent\_index + \beta_4 emotiveness + \beta_5 solve + \beta_6 controls + \varepsilon \tag{2}$$

Regression (1) takes into account the number of reviews posted (*no_reviews*) as dependent variable, whereas regression (2) considers the logarithm of the Amazon user rank as dependent variable (*ln_rank*). As a consequence, regression (1) focuses on explaining the quantity of user generated content contributed. In contrast, since regression (2) has a focus on the Amazon user rank, also the quality of the contributions is considered. The regressions are outlined in equation 1 and 2.

Empirical Study

Descriptive Statistics

In total, the dataset acquisition resulted in a sample of 30,707 users whose profile data was extracted from their individual user profiles publicly available on amazon.com. Table 2 presents the descriptive statistics of this sample. Focusing on the number of reviews (*no_reviews*) posted by the users analyzed, each user has posted on average 22 reviews, whereas the user with the highest number of reviews has contributed 7,815 reviews. As follows, our sample covers a wide range of contributors. The helpfulness ratings of these users' reviews range from 0.012 to 1, whereas on average, the reviews are perceived as helpful by 72.8% of the voters. Thus, the average review is seen as rather useful within the purchase decision process. If the amount of users taking part in the Amazon vine program is taken into account, we find that only 2.2% of the users in the sample are given the economic incentive to contribute reviews about products that are provided for free. Furthermore, 72.6% of the users analyzed are male.

Table 2. Descriptive Statistics (Stdev. = Standard Deviation)				
	Min	Mean	Max	Stdev.
no_reviews	1	22.425	7815	118.384
ln_rank	0.693	13.787	16.271	2.014
helpful_percent	0.012	0.728	1	0.244
vine_voice	0	0.022	1	0.148
male	0	0.726	1	0.446
contact_index	0	0.162	1	0.180
interest_index	0	0.151	1	0.174
referent_index	0	0.578	1	0.494
emotiveness	0	0.525	50	2.488
solve	0	0.100	22.222	0.701

Taking into account the information disclosed in the different profiles, the *contact_index,* which covers information that could be used to contact the different users, as well as the *interest_index* that encompasses whether a user has disclosed information about her or his interests, have mean values of 0.162 and 0.151. Thus, on average, users reveal between 15% to 16% of all information that could be disclosed. However, there are also users disclosing the whole information which is indicated by the

maximum of 1 related to the three indices. If *emotiveness* and *solve* are considered, it can be observed that the amount of words related to emotions utilized by the users is five times higher than the amount of words related to problem solving. This is indicated by *emotiveness* of 0.525 and *solve* of 0.100.

Table 3 shows the correlations of the different variables. It can be observed that the correlation between the number of reviews posted as well as a user's rank is negative (-0.38). This is as expected since users posting a lot of reviews might also achieve a better (i.e. lower) rank. Interestingly, there is no high correlation between both variables which indicates that a user posting a large number of reviews is not necessarily ranked highly. Thus, these users' product reviews are not more helpful than other users' reviews. Another reason could be that these users published their reviews during a longer time period since the Amazon ranking also takes into account the time having passed since a review has been posted.

The fact that we find a correlation of 0.42 between the *vine_voice* dummy variable as well as *no_reviews* indicates that Amazon prefers users providing several product reviews or being ranked highly when selecting users for the vine program. Nevertheless, the correlation is not high so that the variables are used within the subsequent analyses. Taking the remaining correlations into account, it is interesting to observe that there are only moderate correlations between the different information disclosure indices of 0.24 to 0.40. Furthermore, *emotiveness* and *solve* are only moderately correlated (0.33).

		1	2	3	4	5	6	7	8	9	10
Table 3. Variable Correlations											
1	no_reviews	1.00									
2	ln_rank	-0.38	1.00								
3	helpful_percent	0.03	-0.42	1.00							
4	vine_voice	0.42	-0.41	0.06	1.00						
5	male	0.00	-0.02	-0.05	-0.01	1.00					
6	contact_index	0.16	-0.17	0.01	0.20	0.04	1.00				
7	interest_index	0.25	-0.32	0.03	0.29	-0.01	0.40	1.00			
8	referent_index	0.10	-0.17	0.03	0.11	0.00	0.39	0.24	1.00		
9	emotiveness	0.18	-0.21	0.02	0.20	-0.02	0.22	0.54	0.14	1.00	
10	solve	0.16	-0.16	0.02	0.16	-0.01	0.17	0.36	0.10	0.33	1.00

Validation of the Research Model

In order to evaluate the research model, we perform two OLS regressions with (1) the number of reviews contributed as well as (2) the logarithm of the Amazon user rank as dependent variables. Standard errors robust against heteroscedasticity are used. Consequently, we analyze the quantity of online product reviews contributed (1) as well as the Amazon user rank (2), which takes into account both the quantity of online product reviews contributed and the quality of those product reviews (Amazon 2013b). Furthermore, taking into account the user rank is useful for validating the results since the user rank also considers temporal aspects of the contribution, whereas recent reviews are weighted higher than past reviews.

Considering regression (1), we can accept research hypothesis H1: the level of information disclosure has a positive impact on the number of online product reviews posted. To be more specific, we find that the contact index (*contact_index*) has a positive influence that is significant at the 5% level of significance. For the interest and referent indices (*interest_index, referent_index*), we find a positive influence that is significant at the 1% level. As follows, the more information disclosed within a user profile, the higher the contribution to user generated content platforms. The *interest_index* has the highest coefficient (61.828) which shows that disclosing the related information has the highest impact on the level of contribution.

Focusing on research hypothesis H2, we confirm a positive relationship between a user's level of emotiveness (*emotiveness*) as well as the contributions to user generated content platforms. This

relationship is significant at the 5% level. Thus, users expressing a greater amount of emotions within the free text of their profile also contribute more product reviews. In addition, we also confirm hypothesis H3, i.e. an increased problem-orientation (*solve*) leads to increased contributions to user generated content platforms in the social commerce context, which is significant at the 1% level. However, taking into account the coefficients of the corresponding variables, the effect size of a user profile containing only emotional or problem-solving related words (which would cause the corresponding variable to be 1) is much smaller compared to the impact of a user disclosing related profile information as measured by the *contact_index* and *interest_index*.

Taking into account the control variables, we confirm that social feedback in form of helpfulness votes received from other consumers for the online product reviews published (*helpful_percent*) is positively related to the number of reviews contributed. This relationship is significant at the 1% level. Moreover, we also confirm that economic incentives in form of the Amazon vine program (*vine_*voice) have a positive relationship to the number of online product reviews published, which is also significant at the 1% level. In this context, the coefficient of 300.385 implies that users taking part in the vine program publish about 300 online product reviews more than regular users. Finally, we find that the gender (*male*) of the user has an influence on the number of online product reviews published, whereas men post 2.295 online product reviews more when compared to women. This relationship is significant at the 10% level.

If the impact of the independent variables on the Amazon user rank is taken into account in regression (2), our results are nearly identical, with the exception of the *contact_index*. In this case, we find no significant influence on the Amazon user rank. Thus, H1 can only be partly accepted in this case. In contrast, the influence of the other independent variables is significant at the 1% level. Furthermore, the signs of the coefficients estimated in regression (2) are reversed when compared to regression (1). This is caused by the definition of *ln_rank*, since the lower the rank, the better.

		(1) no_reviews		(2) ln_rank	
Table 4. Regression Analysis (n=30,707 user) *** p<10%, ** p<5%, *** p<1% (robust standard errors)**					
		Coefficient	p-value	Coefficient	p-value
	constant	-5.021	0.005***	16.895	0.000***
C1	helpful_percent	3.759	0.003***	-3.287	0.000***
C2	vine_voice	300.385	0.000***	-4.368	0.000***
C3	male	2.295	0.055*	-0.168	0.000***
H1	contact_index	15.739	0.020**	0.057	0.381
	interest_index	61.828	0.000***	-2.101	0.000***
	referent_index	5.129	0.000***	-0.296	0.000***
H2	emotiveness	1.144	0.017**	-0.0217	0.000***
H3	solve	9.195	0.000***	-0.067	0.000***
	F-Value	76.22	0.000***	2354.70	0.000***
	Adjusted R^2	0.200		0.380	

The adjusted R^2 of 0.200 for regression (1) shows that 20% of the variance of the number of reviews posted are explained by our research model. In contrast, if regression (2) is taken into account, 38% of the variance of *ln_rank* is explained. Thus, the different independent variables of our study explain an acceptable amount of the variance. Since the Amazon user rank directly takes into account the helpfulness of a user's online product reviews weighted by the time the online product reviews are published, it is understandable why the adjusted R^2 of regression (2) is comparatively higher than in the case of regression (1). If regression (2) is re-estimated by also taking into account *no_reviews* as independent variable, the results also remain robust.

Considering the goodness of our results, we find that the null hypothesis that every coefficient is zero can be rejected at the 1% level of significance, which is indicated by the *F*-Values of 76.22 (regression 1) as well 2354.70 (regression 2). Furthermore, to test for multicollinearity, we calculate variance inflation factors for each independent variable. In each case, no multicollinearity was detected since the highest score of 1.72 is below common thresholds of 4 or 10 (O'Brien 2007).

Discussion

The proposed research model integrates signaling theory, social information processing theory and communication theory to explain which user characteristics explain the contribution of user generated content. Our results reveal that user characteristics extracted from user profiles help to explain the contribution of user generated content in the social commerce context. We observe an influence of the amount of information disclosed, emotiveness as well as problem-orientation. Furthermore, we confirm that social feedback, economic incentives as well as gender have an influence on the level of contribution.

The information disclosed in user profiles is positively related to the contribution of user generated content, which supports the argumentation that in both cases, users providing such information send signals in order to be perceived as a more reliable part of the community and to increase social interactions. Nevertheless, the *contact_index* has no significant influence if the Amazon user rank is taken into account, which indicates that signaling interests and points of reference are more effective indicators for the amount of contribution in this case. One possible explanation is that users mostly focusing on communication contribute a large number of product reviews, as shown by regression (1). However, these users do not necessarily contribute helpful product reviews, as shown by the non-significant coefficient in regression (2). In addition, since emotiveness has a positive impact as well, we conclude that emotional users also contribute an increased amount of online product reviews. Finally, users being more problem-oriented and thus pursuing more supportive communication are also shown to contribute an increased amount of user generated content.

From a methodological perspective, we confirm that analyzing the text provided in user profiles contains valuable information. Although the texts provided by the users are rather short, this shows that online retailers should offer their users the possibility to disclose such information. On the one hand, their users can send signals to other consumers, and on the other hand, online retailers are then able to analyze this information.

We are aware that the time a user has registered and consequently, the time span a user is active on a social commerce website might also have an influence on the amount of reviews posted. However, in case of the number of total online product reviews posted, controlling for this time span is not possible based on the dataset acquired from Amazon since the user profiles do not contain any data field referring to the time of registration. Fortunately, the Amazon user ranking takes into account the time when reviews were posted since recent ones are given more relevance than older ones (Amazon 2013b). When analyzing the Amazon user rank in comparison with the total number of postings, our results are robust, except from the diminishing impact of the contact index. Thus, we conclude that also the results in the first regression taking into account the total number of postings are reliable, i.e. possible time effects cancel out due to the large number of users analyzed. Furthermore, as the results of the second model show, the different factors taken into account in this study can also explain the contribution behavior in terms of review quantity and quality, so that next to identifying users contributing a large quantity of online product reviews, also users contributing high-quality reviews can be identified.

In our research model, we also control for economic incentives as well as social feedback operationalized by means of a dummy variable measuring if a user takes part in the Amazon vine program as well as by means of a variable measuring a user's helpfulness assessments received. In this case, for both variables, it can be questioned whether a causality related to the contribution of user generated content can be assumed. On the one hand, Amazon vine memberships are given to users on the basis of whether they provided helpful reviews in the past (Amazon 2013a). On the other hand, it can also be assumed that if a user writes many product reviews, he gets more experience on how to write such reviews which causes his reviews to be more helpful. In both cases, a reversed causality could also be possible. Since our results could be biased if this reversed causality would be assumed, we also run our regressions without taking the control variables into account. In this case, the results of our study remain robust.

In recent years, a discussion about privacy issues in the context of disclosing profile information within the web and especially in social media has started (Acquisti and Gross 2006). In this respect, the willingness of users to provide information on their user profiles might change and thus, it might reduce the possibility for online retailers to identify the most valuable users. Consequently, online retailers should implement privacy mechanisms, for instance enabling users to share their profile information only with other registered users, in order to keep the advantage of being able to analyze information disclosed by their users. Furthermore, in order to be able to obtain a comprehensive view of their users, online retailers might take pre-defined user profile ontologies into account that propose several classes of profile fields that could be included in the user profiles (Golemati et al. 2007).

Another important motivation for users to contribute online product reviews is altruism, i.e. helping others without expecting a reward in return (Cheung and Lee 2012). Due to the variable operationalization based on the information disclosed in user profiles, a direct measure of altruism as in the case of a survey is not possible. However, the concept of problem-orientation is closely related to this topic since it covers the aspect of being orientated to problem solving. Thus, although not measured directly, altruism is covered indirectly within the research model. Finally, we are also aware that our current study only covers contributors posting online product reviews on amazon.com. In this context, it can be assumed that the results are relevant for other online retailers providing similar services. We expect amazon.com to be a representative data source since this company offers a variety of products and thus, the risk that the results are biased because of a specific user group preferring only a certain product category is reduced. This is also evidenced by the fact that, in total, the users analyzed within this study contributed 688,618 reviews about a variety of products.

Conclusion

Online product reviews have become an important asset for online retailers since online consumers take them into account when making their purchase decisions. Thus, providing the opportunity to write and read online product reviews can increase the popularity of an online retailer's website and consequently, increase sales and profits. Online retailers might foster the contribution of user generated content by means of providing incentives to the users being most promising to add online product reviews. However, little is known of what user characteristics are related to the amount of content contributed since previous research has mainly concentrated on the question of what motivates users to post user generated content and has neglected information disclosed in online profiles.

With our study, we contribute to the literature on user generated content in social commerce regarding the research question of which user characteristics are related to the contribution of user generated content in form of online product reviews. We therefore build upon signaling theory, social information processing theory, communication theory as well as motivational aspects identified in previous research. We develop and evaluate a research model to explain the contribution to user generated content platforms in the social commerce context. We enhance the previous understanding by finding that the level of information disclosure, the level of emotiveness as well as the level of problem-orientation positively influence the level of contribution, measured either as the total number of reviews or by means of the Amazon user rank. In this context, the user rank also takes into account review helpfulness as well as temporal aspects. We confirm the impact of social feedback, economic incentives as well as gender. Furthermore, we contribute to the literature by proposing a methodology to identify the most active contributors by analyzing user profiles including free texts. In contrast to previous survey-based studies, this has the advantage that a large number of users can be analyzed and that no bias related to certain groups of users exists that (do not) participate in such a survey.

From a practical point of view, our study is relevant due to multiple reasons. At first, we provide an understanding for online retailers of which factors are important to identify consumers that potentially contribute online product reviews. Thus, online retailers might analyze the user profiles of newly registered consumers based on their level of information disclosure as well as on emotions and problem-orientation expressed. Then, further incentives for those users potentially contributing valuable online product reviews could be provided. In this context, we also confirm that economic incentives contribute to content provision, in this case in form of the Amazon vine program that offers consumers products free of charge in exchange for reviewing those. Since next to economically related incentives, social feedback is also an important driver for providing online consumer reviews, online retailers might offer increased

possibilities for reviewers to obtain social feedback. Next to general helpfulness votes, it might also be possible to present statistics to online consumers indicating which of their reviews are popular, i.e. displayed frequently. Furthermore, the importance of different user profiles shows that online retailers should offer their consumers the opportunity to maintain user profiles at all.

Within this study, we focus on the aspects influencing the level of contribution to user generated content platforms within the social commerce context. In order to provide further decision support for online retailers on how to identify the most important contributors, we plan to take our results into account and also evaluate different machine learning models in order to investigate which technology might be ideally applied within the decision support system context. Furthermore, the current study takes into account users posting reviews on amazon.com. Since it is possible that the user characteristics influencing the contribution to user generated content platforms might differ across cultures, one might further repeat this study for different online retailers based in other countries in order to examine potential differences. Finally, previous research in the context of review diagnosticity theory has found that product reviews differ, for instance related to review length or star rating depending on whether they are dealing with search goods or experience goods. Thus, we plan to evaluate whether the users publishing mainly product reviews related to search goods differ from users publishing reviews related to experience goods, and vice versa.

References

Acquisti, A., and Gross, R. 2006. "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook," in *Privacy Enhancing Technologies*, G. Danezis, and P. Golle (eds.): Springer Berlin Heidelberg, pp. 36-58.

Amazon 2013a. "Amazon Vine Program - What is Amazon Vine?" *http://www.amazon.com/gp/vine/ help (accessed: 05/02/2013)* .

Amazon 2013b. "Amazon's Top Customer Reviewers - How ranking works," *http://www.amazon.com/ gp/customer-reviews/guidelines/top-reviewers.html (accessed: 05/02/2013)*.

Back, M. D., Schmukle, S. C., and Egloff, B. 2008. "How extraverted is honey.bunny77@hotmail.de? Inferring personality from e-mail addresses," *Journal of Research in Personality* (42:4), pp. 1116–1122.

Boyd, B. K., Dess, G. G., and Rasheed, A. M. A. 1993. "Divergence between Archival and Perceptual Measures of the Environment: Causes and Consequences," *The Academy of Management Review* (18:2), pp. 204–226.

Cheung, C. M. K., and Lee, M. K. O. 2012. "What drives consumers to spread electronic word of mouth in online consumer-opinion platforms," *Decision Support Systems* (53:1), pp. 218–225.

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.

Coley, A., and Burgess, B. 2003. "Gender differences in cognitive and affective impulse buying," *Journal of Fashion Marketing and Management* (7:3), pp. 282–295.

Curty, R. G., and Zhang, P. 2011. "Social commerce: Looking back and forward," *Proceedings of the American Society for Information Science and Technology* (48:1), pp. 1–10.

Dholakia, R. R., and Chiang, K.-P. 2003. "Shoppers in Cyberspace: Are They From Venus or Mars and Does It Matter?" *Consumers in Cyberspace* (13:1–2), pp. 171–176.

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

Fiore, A. T., Taylor, L. S., Mendelsohn, G., and Hearst, M. 2008. "Assessing Attractiveness in Online Dating Profiles," *Proceedings of the Computer/Human Interaction Conference Florence, Italy*.

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.

Gibb, J. R. 1961. "Defensive Communication," *Journal of Communication* (11:3), pp. 141–148.

Gill, A. J., French, R. M., Gergle, D., and Oberlander, J. 2008. "The language of emotion in short blog texts," *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, San Diego, CA, USA* , pp. 299–302.

Golemati, M., Katifori, A., Vassilakis, C., Lepouras, G., and Halatsis, C. 2007. "Creating an Ontology for the User Profile: Method and Applications," *Proceedings of the First International Conference on Research Challenges in Information Science, Ouarzazate, Morocco.*

Gosling, S. D., Gaddis, S., and Vazire, S. 2007. "Personality Impressions Based on Facebook Profiles," *Proceedings of the International Conference on Weblogs and Social Media, Boulder, Colorado, USA.*

Hancock, J. T., Landrigan, C., and Silver, C. 2007. "Expressing Emotion in Text-based Communication," *Proceedings of the Computer/Human Interaction Conference San Jose, California, USA.*

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.

Hernández, B., Jiménez, J., and Martín, J. M. 2011. "Age, gender and income: do they really moderate online shopping behaviour?" *Online Information Review* (35:1), pp. 113–133.

Hui, T.-K., and Wan, D. 2007. "Factors affecting Internet shopping behaviour in Singapore: gender and educational issues," *International Journal of Consumer Studies* (31:3), pp. 310–316.

Izard, C. E. 2002. "Translating Emotion Theory and Research Into Preventive Interventions," *Psychological Bulletin* (128:5), pp. 796–824.

Kollock, P. 1999. "The Economies of Online Cooperation: Gifts and Public Goods in Cyberspace," in *Communities in Cyberspace*, M. Smith, and P. Kollock (eds.), London, UK: Routledge, pp. 220–239.

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.

Krumm, J., Davies, N., and Narayanaswami, C. 2008. "User-Generated Content," *Pervasive Computing, IEEE* (7:4), pp. 10–11.

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.*

Lawler, E. 1984. "The Strategic Design of Reward Systems," in *Strategic Human Resource Management*, C. Fombrun, N. Tichy, and M. Devanna (eds.), New York, NY, USA: Wiley, pp. 127–147.

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.

Ma, M., and Agarwal, R. 2007. "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities," *Information Systems Research* (18:1), pp. 42–67.

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.

Matos, C. A., and Rossi, C. A. V. 2008. "Word-of-mouth communications in marketing: a meta-analytic review of the antecedents and moderators," *Journal of the Academy of Marketing Science* (36:4), pp. 578-596.

Mortensen, D. C. 2008. *Communication Theory*, New Brunswick, New Jersey, USA: Transcation Publishers.

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.

Neuendorf, K. A. 2002. *The Content Analysis Guidebook*, Thousand Oaks, California, USA: Sage Publications.

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.

O'Brien, R. M. 2007. "A Caution Regarding Rules of Thumb for Variance Inflation Factors," *Quality & Quantity* (41), pp. 673–690.

Rosenberg, S. D., Schnurr, P. P., and Oxman, T. E. 1990. "Content Analysis: A Comparison of Manual and Computerized Systems," *Journal of Personality Assessment* (54:1&2), pp. 298–310.

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

Smith, D. H. 1981. "Altruism, Volunteers, and Volunteerism," *Nonprofit and Voluntary Sector Quarterly* (10:1), pp. 21–36.

Social Security Administration 2011. "Popular Baby Names -Beyond the Top 1000 Names," *http://www.socialsecurity.gov/OACT/babynames/limits.html (accessed: 05/02/2013)*.

Sonnier, G. P., McAlister, L., and Rutz, O. J. 2011. "A Dynamic Model of the Effect of Online Communications on Firm Sales," *Marketing Science, in Press.*

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

Sprague, R. H. 1980. "A Framework for the Development of Decision Support Systems," *MIS Quarterly* (4:4), pp. 1–26.

Stoeckl, R., Rohrmeier, P., and Hess, T. 2007. "Motivations to produce User Generated Content: differences between webloggers and videobloggers," *Proceedings of the 20th Bled eConference, Bled, Slovenia.*

Stone, P. J., Bales, R. F., Namenwirth, J. Z., and Ogilvie, D. M. 1962. "The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information," *Behavioral Science* (7:4), pp. 484–498.

Tong, Y., Wang, X., and Teo, H.-H. 2007. "Understanding the Intention of Information Contribution to Online Feedback Systems from Social Exchange and Motivation Crowding Perspectives," *Proceedings of the 40th Hawaii International Conference on System Sciences, Hawaii, USA.*

Utz, S. 2009. "'Egoboo' vs. altruism: the role of reputation in online consumer communities," *New Media & Society* (11:3), pp. 357–374.

van Slyke, C., Comunale, C. L., and Belanger, F. 2002. "Gender Differences in Perceptions of Web-Based Shopping," *Communications of the ACM* (45:7), pp. 82–86.

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.

Walther, J. B. 1992. "Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective," *Communication Research* (19:1), pp. 52–90.

Walther, J. B., Loh, T., and Granka, L. 2005. "Let Me Count the Ways: The Interchange of Verbal and Nonverbal Cues in Computer-Mediated and Face-to-Face Affinity," *Journal of Language and Social Psychology* (24:1), pp. 36–65.

Wang, J., Ghose, A., and Ipeirotis, P. G. 2012. "Bonus, Disclosure, and Choice: What Motivates The Creation of High-Quality Paid Reviews?" *ICIS 2012 Proceedings*.

Wasko, M. M., and Faraj, S. 2005. "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly* (29:1), pp. 35–57.

Weber, R. P. 1984. "Computer-aided content analysis: A short primer," *Qualitative Sociology* (7:1), pp. 126–147.

Weber, R. P. 1990. *Basic Content Analysis*, Newbury Park, California, USA: Sage.

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*.

Yin, D., Bond, S., and Zhang, H. 2011. "Dreading and Ranting: The Distinct Effects of Anxiety and Anger in Online Seller Reviews," *ICIS 2011 Proceedings.*

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.