ASSESSING QUALITY OF CONSUMER REVIEWS IN MOBILE APPLICATION MARKETS: A PRINCIPAL COMPONENT ANALYSIS APPROACH

Sangwook Ha, Department of Information Systems, National University of Singapore, Singapore, sangwook@comp.nus.edu.sg

Abstract

This study presents a simple, theory-based method for calculating a metric which reflects the quality of online consumer reviews in mobile application markets. Derived from prior online consumer review studies based on psychology, information quality, and economics literature, a metric for measuring online consumer review quality is developed. The metric is a weighted sum of three variables (Squared Star Rating, Log-transformed Word Count, and Sum of Squared Negative and Positive Sentiment), and weights for calculating the metric are estimated by using Principal Component Analysis (PCA) technique. Preliminary assessment of the proposed method shows that metrics computed by using the proposed method are positively correlated with helpfulness ranks of mobile application reviews in Google Play. However, PCA results show that one of the variables (i.e., sentiment) used for developing the metric did not load consistently on the first factor component. From the findings of the preliminary evaluation on the metric, limitations and future research directions of the proposed method are discussed.

Keywords: Mobile Application Market, Online Consumer Review, Review Sentiment.

1 INTRODUCTION

As the smartphone market grows, the number of mobile applications for smartphones is expanding rapidly. For example, Apple reported that people spent over $10 billion on its iTunes App Store in 2013 (Apple Inc. 2014) and Google's CFO noted that Google Play, a mobile application store for its 'Android' platform smartphones drove its recent revenue growth (Popper 2014). According to Gartner, revenues from mobile application markets will hit $77 billion in 2017 (Gartner 2014a). This trend has not only opened up new mobile business opportunities but also created novel research opportunities on mobile businesses (Ladd et al. 2010; Scornavacca et al. 2006).

Among various research topics in this area, one prominent question is "how to help potential customers to discover applications which meet their needs?" People are aware of only a fraction of applications in the vast mobile application markets (Gartner 2014b). Moreover, numerous "me-too" (e.g., Flappy Bird and its countless imitators) and malicious applications (e.g., fake or phishing mobile applications) hinder consumers to find applications which they really want. To address this issue, researchers have proposed several solutions, such as a search engine for mobile application markets (Datta et al. 2013), and a recommender system for mobile applications (Yan & Chen 2011). Although these systems have demonstrated their utility to a certain extent in real-world situations, researchers are still looking for additional methods to improve them as the volume of the market is expanding fast (Yan & Chen 2011).

An emerging idea for promoting and presenting various kinds of products to potential customers effectively is using online consumer reviews (Clemons et al. 2006). The impact of online consumer reviews of a product on its sales is a well-known phenomenon (Chevalier & Mayzlin 2006; Duan et al. 2008a) and studies posit that displaying 'helpful reviews' (i.e., high-quality consumer reviews) could strengthen this relationship (Ghose & Ipeirotis 2011; Mudambi & Schuff 2010) because information from such peer-reviews could assist consumers' decision-making processes (Sun 2012).

In the past, researchers attempted to address this idea by proposing algorithms and building systems for automatically assessing 'helpfulness' of each review based on data mining techniques and learning algorithms in the context of online marketplaces like Amazon.com (e.g., Cao et al. 2011; Chen & Tseng 2011; Kim et al. 2006; Lee & Choeh 2014; Liu et al. 2007; Tsur & Rappoport 2009). However, applying these approaches directly to mobile application markets are sometimes challenging because 1) some features of online consumer reviews required for implementing previous methods are often system-dependent - e.g., most mobile application markets do not provide detailed information of the author of a review, which is one of the crucial factors for measuring the quality of consumer reviews in previously developed methods, 2) prior approaches require numbers of pre- and post-processing steps for getting results (e.g., need to train algorithms or build lexicons before using them) so it is questionable that these methods can effectively meet the demands of the fast-growing mobile application markets 3) prior studies mainly focus on predicting helpfulness of an online consumer review (e.g., numbers of helpfulness votes or a proportion of helpfulness votes in the number of total votes an online consumer review received) rather than providing quality index or score of each online consumer review which could be more helpful for practitioners and consumers (one exception is RevRank, Tsur & Rappoport 2009), and 4) while prior studies built upon a data mining approach often assume a linear relationship between features of an online consumer review and helpfulness of the online consumer review, recent empirical studies find non-linear relationships between them, casting doubts on this assumption (e.g., Baek et al. 2012; Mudambi & Schuff 2010).

To mitigate these issues in the extant research, this study presents a simple, theory-based method for calculating a metric which reflects quality of consumer reviews in mobile application markets. Developed from online consumer review studies based on psychology, information quality, and economics literature (e.g., Baek et al. 2012; Ghose & Ipeirotis 2011; Mudambi & Schuff 2010; Otterbacher 2009), a metric for assessing quality of online consumer reviews is devised. The metric is a weighted sum of three variables (i.e., Squared Star Rating, Log-transformed Word Count, and Sum of Squared Negative and Positive Sentiment), and the first factor component of Principal Component Analysis (PCA) on these three variables is used as weights. Preliminary assessment of the metric shows that it is positively correlated with ranks of consumer reviews of mobile applications in Google Play. Yet one of the variables (i.e., sentiment) did not load consistently on the first factor component (i.e., weights for calculating the metric) in PCA results.

This study contributes to both research and practice. In terms of research contribution, it attempts to synthesize two streams of research on online consumer reviews – empirical studies on factors affecting helpfulness of online consumer reviews and online consumer review recommendation systems – and extends them into the new, emerging area of mobile application markets. For practice, it presents a simple, theory-based metric for choosing high-quality consumer reviews in mobile application markets.

2 LITERATURE REVIEW

Derived from prior literature, this study defines online consumer reviews as peer-generated post-purchase evaluations of product, service, and seller provided by a focal or third-party service (Chen & Xie 2008; Mudambi & Schuff 2010; Pavlou & Dimoka 2006). The format of most online consumer reviews is comprised of three components: 1) star rating (a consumer's subjective score of a product), 2) author profile (information of an author of the review; ranged from a simple user name to a detailed user history depending on the online consumer review system), and 3) review comment (an open-ended written comment on the product). Generally, people who purchased a product or a service on an online market which has its own consumer review system (e.g., Amazon.com, iTunes Store, and Google Play) could write a review of it. In addition, consumers can write their purchase experiences on third-party consumer review services (e.g., metacritic and Rotten Tomatoes).

<table>
<tr><th colspan="2">Variable</th><th>Relationship with Review Quality</th><th>Issues Raised</th><th>Major Reference</th></tr>
<tr><td colspan="2">Star Rating</td><td>Quadratic relationship with review quality</td><td>May vary by types of products</td><td>Mudambi & Schuff (2010); Yin et al. (2014)</td></tr>
<tr><td colspan="2">Author Profile</td><td>Reviews from high-ranked author positively affect review quality</td><td>Inconsistent findings
Information is not always available; depend on online markets/services</td><td>Ghose & Ipeirotis (2011); Otterbacher (2009)</td></tr>
<tr><td rowspan="3">Review Comment</td><td>Word Counts</td><td>Linear and log linear relationship with review quality</td><td>Log linear model showed larger coefficients than linear models</td><td>Baek et al. (2012); Korfiatis et al. (2012); Mudambi & Schuff (2010)</td></tr>
<tr><td>Sentiment</td><td>'Extreme' sentiments contribute to review quality</td><td>No standard method for measuring sentiments</td><td>Baek et al. (2012); Ghose & Ipeirotis (2011); Willemsen et al. (2011)</td></tr>
<tr><td>Readability</td><td>Linear relationship of readability index with review quality</td><td>Inconsistent findings; vary by types of deployed readability index</td><td>Ghose & Ipeirotis (2011); Korfiatis et al. (2012)</td></tr>
</table>

Table 1. A Summary of the Literature Reviewed

Prior research on online consumer reviews investigated reviews-sales relationships and revealed a positive link between them in various online transaction contexts. For example, the volume of online consumer reviews (Chevalier & Mayzlin 2006; Duan et al. 2008a, 2008b) and dispersion of star ratings (Clemons et al. 2006) regarding a product were formed to be positively related to its sales. However, inconsistent findings between studies from this area (e.g., the impact of volume, valence, and dispersion of consumer reviews on sales is inconsistent; Dellarocas et al. 2007) drove researchers to study this topic from different perspectives. One of the alternative approaches, the quality of the online consumer reviews (often referred as 'helpfulness') is recently gaining research attention because providing high-quality reviews to potential customers helps their purchase decision-making process and thus can facilitate product sales (Ghose & Ipeirotis 2011; Mudambi & Schuff 2010). The following section presents the relationship between three components in an online consumer review (star rating, author profile, review comment) and the quality of the review based on our analysis of the previous literature (See Table 1).

2.1 Star Rating

A quadratic relationship between an online consumer review's star rating and its quality is suggested by both theory and empirical results. In marketing, the theory of two-sided advertising argues that providing not only positive but also negative information of a product could enhance consumers' credibility of the information (Crowley & Hoyer 1994). Mudambi & Schuff (2010) elaborated this assertion with prior literature on star rating and proposed a concept of "review extremity" which posits that the squared star rating in an online consumer review influences the quality of the review. Although their initial hypotheses suggest this relationship depends on the type of a product (e.g., a positive relationship with search goods and a negative relationship with experience goods), results from subsequent studies postulate that this relationship could exist regardless of the type of products (Baek et al. 2012; Korfiatis et al. 2012; Yin et al. 2014). In addition, inconsistent findings from studies which tested a linear relationship (Pan & Zhang 2011; Zhang et al. 2010) also suggest the possibility of a quadratic relationship between star rating and review quality.

2.2 Author Profile

Studies showed a review reader's perception of the credibility and expertise of the author of an online consumer review could influence the quality of the review (Baek et al. 2012; Ghose & Ipeirotis 2011; Otterbacher 2009; Schlosser 2011). Psychological models on persuasion (i.e., the impact of external information on a person's decision making process) have found the perceived characteristics of the information sender (e.g., credible, expert) could affect the audiences' decision on a certain topic (Chaiken 1980; Petty & Cacioppo 1986). Inspired by these models, researchers hypothesized that if the author of an online consumer review discloses his/her real name in the review and has a reputation in an e-commerce site (e.g., high-ranked reviewer in Amazon.com), readers could perceive it "better" quality than reviews from anonymous, low-reputation authors. However, findings were inconsistent among studies. For example, disclosing the author's real name in an online consumer review did not statistically affect on the quality of the review in some studies (Baek et al. 2012; Otterbacher 2009). Similarly, the association between an online consumer review author's "reviewer ranking" in Amazon.com and the quality of his/her review was marginal or not clear (Ghose & Ipeirotis 2011; Otterbacher 2009). In addition, "author profile" is a website dependent feature. Even nowadays, only a few e-commerce websites (e.g., Amazon.com) have this feature and most mobile application markets do not have this feature in their system for now. This issue limits further research on this topic and prevents its application in the context of this study (i.e., mobile application markets).

2.3 Review Comment

Comments in online consumer reviews have been studied in several ways. First, researchers focused on the amount of information in an online consumer review. Cognitive psychology and economic literature posit that the availability of information biases a person's decision-making process (Tversky & Kahneman 1974) but it could be bounded because its utility could decrease with consumption (Gossen 1983) and human's limited processing capacity (Lang 2000) could constrain the amounts of external information processed. Based on these arguments, researchers used word counts of an online consumer review as a proxy for the amount of information and tested its relationship with quality of the review. Results support both linear and log linear relationships (Baek et al. 2012; Korfiatis et al. 2012; Mudambi & Schuff 2010; Otterbacher 2009) but the regression coefficient from a log linear model indicates larger effect size and higher statistical significance (Baek et al. 2012) than linear models.

Second, research found there is a link between an online review's sentiments (e.g., subjectivity, polarity) and the quality of the review. Negativity and positivity biases (Herr et al. 1991; Ito et al. 1998) are well-known human traits. Thus, researchers conducted lab experiments (Gershoff et al. 2003; Schlosser 2011) and regression analyses (Baek et al. 2012; Ghose & Ipeirotis 2011; Willemsen et al. 2011) to discover if these effects also exist in the online consumer review context. Although results from both research methods confirmed the existence of those biases (readers of online consumer reviews prefer either very positive or negative reviews than "moderate" reviews), low regression coefficients (Baek et al. 2012) and marginal support for the negativity and positivity bias hypotheses (Ghose & Ipeirotis 2011; Willemsen et al. 2011) suggest that further studies are needed to refine measures for capturing sentiments in online consumer reviews.

Last, some studies investigate the "readability" of an online consumer review and its impact on the quality of the review. Researchers argue that reviews which are written in "readable" style could reduce readers' cognitive efforts and thus they could favor the review more than others with language errors and complex sentence structures (Ghose & Ipeirotis 2011; Korfiatis et al. 2012). However, their tests revealed inconsistent relationships among different types of ease-of-reading indices and the quality of online consumer reviews.

3 PROPOSED METHOD

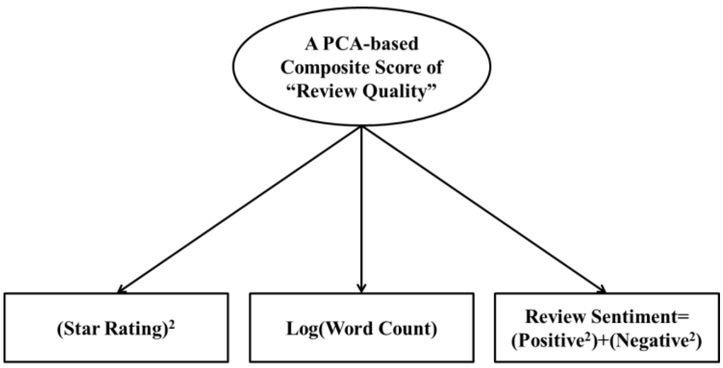

Figure 1. A Graphical Representation of the Proposed Method

Based on the literature review in section 2, a metric for assessing the quality of online consumer reviews in mobile application markets is proposed in Figure 1 and Equation (1). As an equation form, the metric of review quality for the k^{th} review from a group of reviews j (or application j) is a weighted sum of three variables of the k^{th} review (i.e., Star Rating_{jk}^2, Log(Word Count_{jk}), and Review Sentiment_{jk}). Weights for variables from a group of reviews j (or application j; β_{0j}, β_{1j}, and β_{2j}) are acquired from the first factor component of PCA on the correlation matrix of three variables from a group of reviews j (or application j).

$$\textit{Composite Score}_{jk} = \beta_{0j}(\textit{Star Rating}_{jk}^{2}) + \beta_{1j}\{\textit{Log}(\textit{Word Count}_{jk})\} + \beta_{2j}(\textit{Review Sentiment}_{jk}) \quad (1)$$

The metric is rooted in two salient findings of the prior research on online consumer reviews: consumers' preferences on 1) "extreme reviews", and 2) "informative reviews." Both lab experiments and empirical analyses suggest consumers prefer reviews with extreme star rating (Baek et al. 2012; Mudambi & Schuff 2010; Yin et al. 2014) and emotional expressions (Gershoff et al. 2003; Schlosser 2011) than "moderate" reviews. In addition, studies also show consumers perceive long review comments as an indicator of a high-quality review (Baek et al. 2012; Korfiatis et al. 2012; Mudambi &

Schuff 2010; Otterbacher 2009) but cognitive theories suggest its impact may be hampered due to human's limited capacity to consume and process external information (Gossen 1983; Lang 2000). Therefore, the method employs squared star rating and sum of squared positive and negative sentiments to reflect review extremity, and uses log transformed word count to represent the amount of information in a review. Then, a composite score of a review's quality is calculated with equation (1) and weights of each variable estimated from PCA on the correlation matrix of three variables. This process is similar to Singular Value Decomposition (SVD) which commonly used in computer science and related research areas for reducing dimensions of data (Gerbrands 1981). This is because the first factor explains the largest amount of variances of the three inputted variables for conducting PCA, and the relationship between the first factor component and the three inputted variables are represented by each variable's factor loading on the first factor component (Suhr 2005). According to its theoretical and empirical foundations mentioned in the literature review section, this method can be validated by investigating two hypotheses below:

Hypothesis 1: Squared Star Rating, Log-transformed Word Count, and Sum of Squared Negative and Positive Sentiment will be positively loaded on the first factor component from PCA.

Hypothesis 2: The composite score will be positively correlated with proxy measures of review quality from mobile application markets

4 PRELIMINARY EVALUATION

4.1 Evaluation Design

In order to evaluate the proposed method and hypotheses, a pilot study was conducted. Google Play was selected as the site for this study due to two reasons. First, it is one of the major mobile application markets in the industry. Second, it provides alternative information (sort reviews by "helpfulness") for inferring the quality of online consumer reviews in each mobile application's download page. Although Google did not disclose its actual mechanism for sorting reviews by helpfulness, the voting system for review comments (helpful/not helpful/spam) suggest that the Google Play's system resembles Amazon.com's helpfulness voting system and thus the order of reviews sorted by helpfulness can be a proxy measure of the quality of each consumer review.

Variable	Mean	Std. Dev.	Min	Max
Star Rating (ra)	3.58	1.64	1.00	5.00
Word Count (wc)	42.37	23.92	8.00	147.00
Positive Sentiment (pos)	0.41	0.24	0.00	0.90
Negative Sentiment (neg)	0.53	0.26	0.00	0.90
Squared Star Rating (ra2)	15.46	10.19	1.00	25.00
Review Sentiment (sen)	0.57	0.18	0.00	0.82
Log(Word Count) (lwc)	3.60	0.54	2.08	4.99
Helpfulness Rank (rnk)	15.50	8.69	1.00	30.00

Table 2. Descriptive Statistics (N=120)

To collect data to test the proposed method, I searched mobile applications in Google Play which frequently updated and rarely contain spam reviews. Based on these criteria, four applications (The Official YouTube app, AndChat, ES File Manager, and Wakelock Detector) were selected. Then I captured first 30 consumer reviews per each application (total 120 reviews) sorted by "helpfulness" which contain ‘meaningful’ review comments (after the 30th comment, most review comments are short, simple words such as “thanks” or “good!”). The data was collected during March-April 2014 period. The dataset contains: 1) star rating, 2) review comment, 3) word count, and 4) helpfulness rank (i.e., the order of consumer review appeared in the application’s page; 1st review = 30, 30th review = 1). To build the metrics for conducting PCA, I used Perkins (Jacob 2010)’s Python NLTK Sentiment Analysis with Text Classification Demo for analyzing review comments’ sentiment. The tool processes inputted chunk of sentences and outputs the portion of positive (e.g., “nice”, “wonderful”) and negative sentiments (e.g., “bad”, “disappointed”) in the inputted chunk of sentences. Stata 13 (StataCorp. 2013) was used for manipulating/testing data for the proposed method. Table 2 shows descriptive statistics of the dataset and Table 3 presents Pearson correlation coefficients between the measures used for the study.

	ra	**wc**	**pos**	**neg**	**ra2**	**sen**	**lwc**	**rnk**
ra	1.00							
wc	0.13	1.00						
pos	0.51*	0.04	1.00					
neg	-0.47*	0.03	-0.50*	1.00				
ra2	0.99*	0.14	0.52*	-0.45*	1.00			
sen	-0.13	0.10	0.34*	0.51*	-0.10	1.00		
lwc	0.10	0.93*	0.03	0.02	0.10	0.08	1.00	
rnk	0.21*	0.33*	-0.02	0.03	0.20*	-0.05	0.37*	1.00

*Table 3. Correlation Matrix (N=120); *p<0.05*

4.2 Preliminary Results

To test the two hypotheses stated in the Section 3, I investigated 1) the factor structure of the first factor component by examining the factor loadings of the first factor component from PCA, and 2) calculated Spearman and Pearson correlations between a metric for each mobile application review calculated by the proposed method and the helpfulness rank of each mobile application review gathered from Google Play. Results show that calculated metrics for mobile application reviews were positively correlated (Pearson and Spearman’s $r=0.31$, $p<0.001$) with their helpfulness ranks (Hypothesis 2, Table 5), but one of the variables (i.e., sen) did not load consistently on the first factor component (Hypothesis 1, Table 4) from PCA.

Post-hoc analyses of each subset (n=30) suggest that factor structure may vary by each mobile application. For example, the sentiment variable (sen) created from reviews on two mobile applications (AndChat and ES File Manager) has a positive loading on the first factor component from PCA (Table 4). However, the sentiment variable (sen) devised from other two mobile applications (YouTube and Wakelock Dector) has a negative loading on the first factor component from PCA.

One possible reason for this inconsistency could be rooted in the characteristics of mobile application. Because the quality of a mobile application can vary by multiple factors (e.g., version of the application, a consumer’s mobile device and its operating system) and it can be enhanced through the application developer’s efforts (i.e., maintaining the application’s code), consumers may ‘report’ issues in the review comments regarding the application rather than ‘express’ their emotions towards their usage experience. An anecdote notice which often comes with a mobile application’s description such as “please report the application’s bugs and issues to the developer’s email rather than user review section” would be an example of this explanation. This result also can be caused by diverse types of sentiment information in review comments which the deployed sentiment analysis method may not be able to assess properly. Recent studies on sentiment analysis find some kinds of sentiment information can change the overall sentiment assessment of an online posting – e.g., non-verbal expressions like emoticon (Hogenboom et al. 2013) can provide additional sentiment information, and sarcasm can reverse ”positive” sentiment into ”negative” sentiment or vice versa (González-Ibáñez et al. 2011; Riloff et al. 2013).

	Pooled (N=120)	**YouTube (N=30)**	**AndChat (N=30)**	**ES File Manager (N=30)**	**Wakelock Detector (N=30)**
ra2	0.7964	0.7639	0.6556	0.6814	0.7365
sen	-0.4448	-0.4697	0.5991	0.2046	-0.3307
lwc	0.4099	0.4316	0.4596	0.7027	0.5901
Variance Extracted	0.3681	0.3822	0.4317	0.5086	0.4909

Table 4. Factor Structure of the 1st Component (Unrotated)

	Pooled (N=120)	**YouTube (N=30)**	**AndChat (N=30)**	**ES File Manager (N=30)**	**Wakelock Detector (N=30)**
Spearman	0.31***	0.49**	0.29	0.30	0.53**
Pearson	0.31***	0.46**	0.28	0.35*	0.48**

*Table 5. Correlation Between Composite Scores and Helpfulness Ranks; ***p<0.001, **p<0.01, *p<0.5, other coefficients are significant between p<0.11-0.15 range*

5 LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Like all other studies, this study has its own limitations. First, the pilot study used relatively small size of data (N=120) from three categories (Music and Video, Communication, and Productivity) for the analysis. Some categories in the Google Play (e.g., games and entertainments) are excluded from the preliminary analysis because 1) many mobile games and entertainment services have an “invitation code” system which forces their users to write ‘promotion reviews’ for receiving rewards (e.g., free items for games, bonus online credits for purchasing services), and 2) spam reviews prevail among those categories. Hence, the generalizability of the study may be limited. Second, as discussed in the section 4.2, some facets of sentiment information may not covered by the implemented method for calculating sentiment in review comments (e.g., emoticon and sarcasm). Third, some studies show that people’s evaluation on online consumer reviews (e.g., helpfulness ranks of reviews) can be biased by social influences and individual traits (Aral 2013; Li & Hitt 2008; Wan & Nakayama 2012) and algorithms for assessing quality of review comments as well (Liu et al. 2007; Tsur & Rappoport 2009). Therefore, further evaluations with different evaluation designs (e.g., field experiments) or comparisons with other metrics for assessing quality of consumer reviews are needed.

These limitations open seveal venues for future works regarding the proposed method and metric: 1) conduct a study on a large set of online review samples from multiple mobile application markets, 2) deploy additional methods for analyzing multiple sentiments sources (e.g., sarcasm and emoticons), 3) assess the quality of the proposed metric with different evaluation methods (and metrics developed for similar purposes).

References

Apple Inc. (2014). App Store Sales Top $10 Billion in 2013. *Apple Inc.*, January 7 (available at http://www.apple.com/pr/library/2014/01/07App-Store-Sales-Top-10-Billion-in-2013.html).

Aral, S. (2013). The Problem With Online Ratings. *MIT Sloan Management Review* 55 (2), 47-52.

Baek, H., Ahn, J., and Choi, Y. (2012). Helpfulness of Online Consumer Reviews: Readers' Objectives and Review Cues. *International Journal of Electronic Commerce* 17 (2), 99–126.

Cao, Q., Duan, W., and Gan, Q. (2011). Exploring determinants of voting for the 'helpfulness' of online user reviews: A text mining approach. *Decision Support Systems* 50 (2), 511–521.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology* 39 (5), 752–766.

Chen, C. C., and Tseng, Y.-D. (2011). Quality evaluation of product reviews using an information quality framework. *Decision Support Systems* 50 (4), 755–768.

Chen, Y., and Xie, J. (2008). Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix. *Management Science* 54 (3), 477–491.

Chevalier, J. A., and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research* 43 (3), 345–354.

Clemons, E., Gao, G., and Hitt, L. (2006). When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry. *Journal of Management Information Systems* 23 (2), 149–171.

Crowley, A. E., and Hoyer, W. D. (1994). An integrative framework for understanding two-sided persuasion. *Journal of Consumer Research* 20 (4), 561–574.

Datta, A., Kajanan, S., and Pervin, N. (2013). A Mobile App Search Engine. *Mobile Networks and Applications* 18 (1), 42–59.

Dellarocas, C., Zhang, X. (Michael), and Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing* 21 (4), 23–45.

Duan, W., Gu, B., and Whinston, A. B. (2008a). Do online reviews matter? -- An empirical investigation of panel data. *Decision Support Systems* 45 (4), 1007–1016.

Duan, W., Gu, B., and Whinston, A. B. (2008b). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing* 84 (2), 233–242.

Gartner. (2014a). Gartner Says by 2017, Mobile Users Will Provide Personalized Data Streams to More Than 100 Apps and Services Every Day. January 22 (available at http://www.gartner.com/newsroom/id/2654115).

Gartner. (2014b). Gartner Says Less Than 0.01 Percent of Consumer Mobile Apps Will Be Considered a Financial Success by Their Developers Through 2018. January 13 (available at http://www.gartner.com/newsroom/id/2648515).

Gerbrands, J. J. (1981). On the relationships between SVD, KLT and PCA. *Pattern Recognition* 14 (1–6), 375–381.

Gershoff, A. D., Mukherjee, A., and Mukhopadhyay, A. (2003). Consumer Acceptance of Online Agent Advice: Extremity and Positivity Effects. *Journal of Consumer Psychology* 13 (1–2), 161–170.

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), 1498–1512.

González-Ibáñez, R., Muresan, S., and Wacholder, N. (2011). Identifying Sarcasm in Twitter: A Closer Look. in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers - Volume 2*, Stroudsburg, PA, USA: Association for Computational Linguistics, 581–586.

Gossen, H. H. (1983). *The laws of human relations and the rules of human action derived therefrom*, Cambridge, Mass.

Herr, P. M., Kardes, F. R., and Kim, J. (1991). Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective. *Journal of Consumer Research* 17 (4), 454–462.

Hogenboom, A., Bal, D., Frasincar, F., Bal, M., de Jong, F., and Kaymak, U. (2013). Exploiting Emoticons in Sentiment Analysis. in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, Presented at the 28th International Symposium on Applied Computing, New York, NY, USA: ACM, 703–710.

Ito, T. A., Larsen, J. T., Kyle, N., and Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology* 75 (4), 887–900.

Jacob, P. (2010). Sentiment Analysis with Python NLTK Text Classification. *Python NLTK Sentiment Analysis with Text Classification Demo* (available at http://text-processing.com/demo/sentiment/).

Kim, S.-M., Pantel, P., Chklovski, T., and Pennacchiotti, M. (2006). Automatically Assessing Review Helpfulness. in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, Presented at the 2006 Conference on Empirical Methods in Natural Language Processing, Stroudsburg, PA, USA: ACL, 423–430.

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), 205–217.

Ladd, D., Datta, A., Sarker, S., and Yu, Y. (2010). Trends in Mobile Computing within the IS Discipline: A Ten-Year Retrospective. *Communications of the Association for Information Systems* 27 (1).

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication* 50 (1), 46–70.

Lee, S., and Choeh, J. Y. (2014). Predicting the helpfulness of online reviews using multilayer perceptron neural networks. *Expert Systems with Applications* 41 (6), 3041–3046.

Liu, J., Cao, Y., Lin, C.-Y., Huang, Y., and Zhou, M. (2007). Low-Quality Product Review Detection in Opinion Summarization. in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Presented at the 2007 Conference on Empirical Methods on Natural Language Processing and Computational Natural Language Learning, Prague, Czech Republic: ACL, 334–342.

Li, X., and Hitt, L. M. (2008). Self-Selection and Information Role of Online Product Reviews. *Information Systems Research* 19 (4), 456–474.

Mudambi, S., and Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *Management Information Systems Quarterly* 34 (1), 185–200.

Otterbacher, J. (2009). Helpfulness' in Online Communities: A Measure of Message Quality. in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Presented at the 27th ACM SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA: ACM, 955–964.

Pan, Y., and Zhang, J. Q. (2011). Born Unequal: A Study of the Helpfulness of User-Generated Product Reviews. *Journal of Retailing* 87 (4), 598–612.

Pavlou, P. A., and Dimoka, A. (2006). The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation. *Information Systems Research* 17 (4), 392–414.

Petty, R. E., and Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*, Springer-Verlag, New York.

Popper, B. (2014). Google Q4 2013: strong revenue growth driven by Play Store and hardware sales. *The Verge*, January 30 (available at http://www.theverge.com/2014/1/30/5362236/google-q4-2013-earnings).

Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N., and Huang, R. (2013). Sarcasm as Contrast between a Positive Sentiment and Negative Situation. in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Presented at the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, Washington, USA: ACL, 704–714.

Schlosser, A. E. (2011). Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments. *Journal of Consumer Psychology* 21 (3), 226–239.

Scornavacca, E., Barnes, S. J., and Huff, S. L. (2006). Mobile Business Research Published in 2000-2004: Emergence, Current Status, and Future Opportunities. *Communications of the Association for Information Systems* 17 (1).

StataCorp. (2013). *Stata Statistical Software: Release 13*, College Station, TX: StataCorp LP.

Suhr, D. D. (2005). Principal component analysis vs. exploratory factor analysis. in *Proceedings of the Thirtieth Annual SAS ® Users Group International Conference*, Presented at the Thirtieth Annual SAS ® Users Group International Conference, Cary, NC: SAS Institute Inc., 203–230.

Sun, M. (2012). How Does the Variance of Product Ratings Matter?. *Management Science* 58 (4), 696–707.

Tsur, O., and Rappoport, A. (2009). RevRank: A Fully Unsupervised Algorithm for Selecting the Most Helpful Book Reviews. in *Proceedings of the Third International Conference on Weblogs and Social Media*, Presented at the Third International AAAI Conference on Weblogs and Social Media, Menlo Park, California: AAAI Press, 154–161.

Tversky, A., and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185 (4157), 1124–1131.

Wan, Y., and Nakayama, M. (2012). Are Amazon.com Online Review Helpfulness Ratings Biased or Not? in *E-Life: Web-Enabled Convergence of Commerce, Work, and Social Life* (M. J. Shaw, D. Zhang, and W. T. Yue. eds.), Springer Berlin Heidelberg, p. 46–54.

Willemsen, L. M., Neijens, P. C., Bronner, F., and de Ridder, J. A. (2011). 'Highly Recommended!' The Content Characteristics and Perceived Usefulness of Online Consumer Reviews. *Journal of Computer-Mediated Communication* 17 (1), 19–38.

Yan, B., and Chen, G. (2011). AppJoy: Personalized Mobile Application Discovery. in *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*, Presented at the 9th international conference on Mobile systems, applications, and services, New York, NY, USA: ACM, 113–126.

Yin, D., Bond, S., and Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly* 38 (2), 539–560.

Zhang, J. Q., Craciun, G., and Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research* 63 (12), 1336–1341.