THREE STUDIES ON EXPLORING THE EFFECTIVENESS OF E-WOM ON CUSTOMER EVALUATION AND BEHAVIOR

LI JING

Ph.D

The Hong Kong Polytechnic University

2018

The Hong Kong Polytechnic University

Department of Management and Marketing

Three Studies on Exploring the Effectiveness of e-WOM on Customer Evaluation and Behavior

Li Jing

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

July 2017

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Li Jing

Abstract

With the advent and prosperity of social media, an increasing number of customers voluntarily post their evaluations and opinions about a product or service online and e-WOM has gained extraordinary growth and aroused intensive academic attention. The current literature has examined how e-WOM, especially in the form of online customer reviews, influences a wide range of outcomes such as customer awareness, perceptions, attitudes, behavioral intentions, and product sales. In my dissertation, I developed three studies that aim to reveal the role of e-WOM in different forms (e.g., photo, text, and vote) on consumer information processing and decision-making.

In the first study, I examined the interaction effects of review certainty, reviewer popularity, reviewer expertise, and the niche width of a restaurant on review usefulness, by drawing on the dual-process theory and social influence theory. Utilizing a zero-inflated negative binomial Poisson regression, I empirically tested my hypotheses based on 10,097 reviews on 2,383 restaurants from Yelp.com. My results indicated that (1) the impact of review certainty on review usefulness decreases with reviewer popularity but does not vary with reviewer expertise; (2) the niche width of a restaurant—as a contextual feature—interacts with review certainty and reviewer characteristics in influencing review usefulness.

In the second study, I focused on one particular type of e-WOM – votes by prior users on a review. I examined how two numeric characteristics of online review helpfulness: 1) helpfulness ratio—the ratio of later viewers who believe that a previous review is helpful; and 2) helpfulness magnitude—the number of later viewers who vote on a previous review; influence consumers' reaction toward the product/service reviewed. Drawing on the social influence theory, this study examined the interactive impacts between these two factors and two other characteristics of online review content (i.e., review valence and type) on consumer trust and attitude. I conducted three lab-based experiments to test the research model. My research finds that regardless of the valence and type of reviews, vote ratio enhances review trustworthiness and guide corresponding product evaluation. In contrast with ratio effect, vote magnitude is significantly influential only for the negative attribute-based review.

In the third study, I investigated three key characteristic of photo e-WOM: layout, sequence and density of photo relative to text in an online customer review. This study conducted both a lab experiment and a field study to demonstrate the role of photo e-WOM on consumer information processing and decision-making. In particular, I adapted the cue summation theory of multi-channel communication to examine the impact of between-channel interactions of text and photo on two outcomes of consumer information processing—diagnosticity and pleasantness of e-WOM and one outcome of consumer decision-making —product value ratings. This study employs both a lab-based experiment and a field analysis to provide robust validation of hypotheses. The experimental results show that separate layout is better than the alternate layout in perceived diagnosticity and product evaluation, especially when the photo is displayed first than the text first displayed. In contrast, for pleasantness, alternate layout is better than separate layout, regardless of the sequence of text and picture. Moreover, the field results suggest that sharing more photos especially outside photos hurts restaurant reputation while sharing photos more on food, drink and menu of a restaurant increases the restaurant's reputation; and for the restaurant as a generalist occupying multiple cuisines, the more photos shared in a review, the better its reputation will be; by contrast, for the restaurant as a specialist occupying few cuisines, the more photos shared on food and drink in a review, the better its reputation will be.

Theoretically, these findings contribute to online customer review literature, provide new managerial implications for leveraging e-WOM, and add new insights into understanding the role of organization positioning in customer evaluations.

Key words: e-WOM; Text e-WOM; Photo e-WOM; Vote e-WOM; Online Voting System; Social Influence Theory; Reviewer Popularity; Reviewer Expertise Certainty; Trust; Social Media

Acknowledgements

I give sincere gratitude to my supervisors, Prof. Eric Ngai and Dr. Xu Xin, for their guidance and support throughout my doctoral study. They are very insightful and considerate of both my academics and my life. Their constructive and patient instructions help me accomplish thesis writing. I am really grateful to have them as my supervisors.

I also thank members of my thesis committee, Prof. SIA Choon Ling, Dr. Xiaojun Zhang, and Dr. Yuwei Jiang, for their advice on my thesis.

Big thanks would go to Dr. Flora Fang who has inspired me to pursue academic work and offered continuous support in my research. I am thankful to spend my doctoral studies with a group of great scholars. I thank the professors and scholars in my PhD program.

I also want to thank the staff who have given me a lot of administrative and research help during my doctoral study, Ms. Candy Lee, Ms. Ann Leung , Ms. Candy Mok, Ms. Parika Woo, and Mr. Ken Chiu. Special thanks to my friends and fellow students for their companionship.

Finally, I want to express my deepest gratitude to my family for their selfless support to my study.

Table of Contents

List of Figures

List of Tables

Introduction

The past decade has witnessed the explosive development and widespread use of social media, accumulating a vast number of user-generated opinions and evaluations for a wide range of products and services in various forms, including text, photos, and videos etc. Customers-generated information is assumed to be more relevant to customers than seller-generated information such as TV commercial advertisement. Further, according to a Nielsen 2013 report[1] on trust in advertising and brand messages, "consumer opinions online" as a form of advertising ranked higher in building customers' trust and facilitating customers to take purchase action, than other traditional advertising forms such as ads on TV, ads in magazines, branded websites, etc. As one important source of customer-generated opinions, e-WOM is articulated as a process in which communications between a sender and a receiver can impact consumer purchase decision process (Cheung et al. 2012), which includes stages of need recognition, information search and processing, alternatives evaluation, purchase decision, purchase and post-purchase assessment (Kotler et al. 2014). While the first e-WOM has existed since the first e-mail, the growth of photo and video sharing, as well as massive user participation exemplified by the voting system, have been striking on online review sites (e.g., Amazon.com, Dine.com).

With the rapid development of technology, many new technics and cues have been created and added to the online review platform, such as the function of uploading photos through which customers not only write out textual comments but also share self-generated photos to other potential customers, and the review voting mechanism via which customers can evaluate previous reviews by clicking associated buttons and the aggregate number of corresponding votes can then be referred to by subsequent customers. As one type of e-WOM, online customer reviews have played an increasingly important role in electronic commerce in various aspects, such as informing potential customers of product knowledge (Martin & Lueg 2013), reducing uncertainty of product quality (Senecal & Nantel 2004), and finally increasing product sales (Chevalier & Mayzlin 2006). The current literature has mainly two specific forms of e-WOM—i.e., textual reviews and numeric ratings. However, online customer reviews consist of not only text and numerical ratings, but also photos and votes. For example, diners at restaurants take photos of food and facilities and post them to rating

[1] http://www.nielsen.com/us/en/insights/reports/2013/the-paid-social-media-advertising-report-2013.html

platforms such as Yelp.com with textual recommendations, which will be voted as helpful or unhelpful by subsequent users via the review helpfulness voting system. In my dissertation, I conduct three studies to explore the effectiveness of these new technics and cues in online review sites. The first study focuses on one specific cue of text review—i.e., certainty, the second study examines the prior users' votes on other user's review, and the third study exploits one important form of e-WOM—i.e., user-generated photos embedded in reviews. The overview of my dissertation is written as below:

Study 1: Effects of Review Content, Reviewer Characteristics, and Organization Niche Width on Review Usefulness

Given the prosperity of online review websites—e.g., Yelp.com—prominently displaying a large scale of online customer reviews; scholars have made efforts to investigate what makes a useful review. However, few insights exist regarding how the review content, reviewer characteristics, and review contexts jointly influence review usefulness. Drawing on the dual-process theory and social influence theory, I examined the interaction effects of review certainty, reviewer popularity, reviewer expertise, and the niche width of a restaurant on review usefulness. Utilizing a zero-inflated negative binomial Poisson regression, I empirically tested my hypotheses based on 10,097 reviews on 2,383 restaurants from Yelp.com. My results indicated that (1) the impact of review certainty on review usefulness decreases with reviewer popularity but does not vary with reviewer expertise; (2) the niche width of a restaurant—as a contextual feature—interacts with review certainty and reviewer characteristics in influencing review usefulness. Theoretically, these findings contribute to online customer review literature and the social media research, provide new guidelines for predicting review usefulness, and add new insights into understanding the role of organization positioning in customer evaluations. In practice, my findings benefit online review platforms (e.g., Yelp.com) to screen and select useful reviews for visitors.

Study 2: How Prior Users' Helpfulness Votes on a Review Influence Subsequent Users' Trust of the Review and Corresponding Product Evaluations in E-commerce Context

With the explosive growth of opinions, news, and product reviews constantly posted on the Internet, massive user participation empowers online opinion evaluation. Opinion evaluation is exemplified by the voting system on third-party review sites (e.g., Amazon.com, Dine.com). I exploit the numeric cues of e-WOM—i.e., user-generated votes on other user' review. Online review platforms provide a review voting system to help customers efficiently figure out or sort out helpful reviews. Such voting mechanism allows consumers to not only view the reviews of predecessors, but also evaluate the reviews by clicking the corresponding buttons. I propose that review votes that are either presented by ratio or magnitude provide users with a direct indicator of the extent to which prior users perceive the review as helpful. I focus on how two numeric characteristics of prior users' votes influence user's attitude towards the product/service reviewed. These characteristics are 1) vote ratio and 2) vote magnitude. The former is the ratio of prior viewers who believe that a review is helpful, and the latter is the total number of prior viewers who vote the review. I draw on social influence theory and propose a two-stage model depicting vote ratio as an unconditional cue and vote magnitude as a conditional cue; the model suggests that users need motivation and ability to exploit magnitude for decision making. I conduct three experiments to test the research model. My research finds that regardless of the valence and type of reviews, vote ratio enhances trustworthiness and guide corresponding product evaluation. In contrast with ratio effect, vote magnitude is significantly influential only for the negative attribute-based review. This finding is attributed to negative reviews, which motivate users to take additional cues for decision making. By contrast, attribute-based review offers users the ability to take a rational reference of prior votes for their product evaluations. These findings make important contributions to review helpfulness literature and extend social influence theory. My findings also provide practical implications for online voting system providers, general participatory sites, and online retailers.

Study 3: A Picture Is Worth One Thousand Words, or Is It? —An Investigation of the Impacts of User-Generated Pictures on Consumer Evaluation in the e-WOM Context

Online social media is flushed with photos with the advancement of digital photography and Web 2.0 technology. Sharing photos has also become popular in the context of electronic Word of Mouth (e-WOM). Nowadays online customer reviews consist of not only textual content and numeric ratings, but also photos. In this study, I focus on one important form of e-WOM—i.e., user-generated photos embedded in reviews. I employ both an experimental study and a field study to provide robust validation of the research model. In experimental study, I adapt the cue summation theory to examine the effects of the relative layout (the alternate layout vs. the separate layout) and relative sequence (photo first vs. text first) on diagnosticity, pleasantness of e-WOM and customers' attitude towards the recommended product/service in the e-WOM. The experimental results show that separate layout is better than the alternate layout in perceived diagnosticity and product evaluation, especially when the photo is displayed first than the text first displayed. In contrast, for pleasantness, alternate layout is better than separate layout, regardless of the sequence of text and picture. In the field study, I examine the interaction effects of photo density and organizational niche width on organizational reputation using the Yelp Challenge Dataset. Utilizing a step-wise regression model, I tested the model using 96,588 photos in 1,241,046 reviews from 15,517 restaurants from Jan 2004 to June 2016, within which 13,995 photos are on the inside, 12,865 photos on the outside, and 15,329 photos on the food, 14,764 photos on the drink, and 15,035 photos on the menu. The empirical results indicate that that (1) sharing more photos especially outside photos hurts restaurant reputation while sharing photos more on food, drink and menu of a restaurant increases its reputation; and (2) for the restaurant as a generalist occupying multiple cuisines, the more photos shared in a review, the better its reputation will be; (3) for the restaurant as a specialist occupying few cuisines, the more photos shared on food and drink in a review, the better its reputation will be.

Chapter 1

Effects of Review Content, Reviewer Characteristics, and Organization Niche Width on Review Usefulness

1.1 Introduction

The contemporary proliferation of online communications by social media has facilitated the generation of a vast number of customer reviews for a wide range of products and services online. As one form of electronic word-of-mouth (e-WOM), online customer reviews, which refer to informational communications among customers concerning the evaluations of goods and services, have played an increasingly important role in electronic commerce in various aspects, such as informing potential customers of product knowledge (Martin & Lueg, 2013; Piccoli, 2016; Safi & Yu, 2017), reducing uncertainty of product quality (Senecal & Nantel, 2004), and increasing product sales (Chevalier & Mayzlin, 2006; Wu & Gaytán, 2013). However, the complete comprehension of a significant number of online customer reviews is a challenging task; hence, the increasing availability of online reviews generates information overload for consumers (Jones *et al.,* 2004). Moreover, customers only consider a small number of reviews that are especially useful for their decision-making. Chen et al. (2008) corroborated that reviews with a high proportion of helpful votes were perceived as of high quality and high ratings—leading to the increase of product sales. Accordingly, designing a mechanism to identify useful reviews has been critically important for practitioners. As a response to this information-overload phenomenon and an answer to the call for identifying useful reviews, social media platforms such as Yelp.com provide a peer voting system that allows the reviewers to grant "useful" votes to a review by asking the question "Was this review…?" under each review. Although this system can identify useful reviews in an ex-post manner, the accumulation of votes needs time and may delay the accessibility of the correct potential customers to right information, which in this case refers to useful reviews. An ex-ante approach to the prediction of review usefulness will help social media platforms screen and select e-WOM appropriately to feed their online visitors with limited time and cognitive resources. My paper intends to offer such an approach to predict the components of a useful review.

At present, product descriptions and slogans generated by sellers frequently use certainty words (e.g., absolutely, must, and always) to persuade potential customers to buy their products or services. In addition, people also hold their opinions with varying degrees of conviction or certainty. For instance, diners who write a review on a new restaurant may differ in the certainty tone, that is, some would be quite sure of their favorable recommendations with the use of certain words such as definitely, completely, absolute, etc. This phenomenon is ubiquitous in online review platforms, and examples are shown in the Appendix A, where the certainty tone is mainly expressed by certain words such as everything, definitely, all, etc. In addition to the ubiquitous existence of certainty tone (e.g., certain, completely, absolutely, sure) in online communications especially product reviews, it also has an effect on persuasion as suggested in the current behavioral literature (e.g., Karmarkar & Tormala 2010; Sniezek & Van Swol 2001). These studies mostly focus on the effects of attitude certainty on persuasion by either relying on self-reports—i.e., how certain/convinced are you of your attitude?, or manipulating degrees of certainty in lab controlled experiment. Although these studies contribute to the identification of the impact of certainty on persuasion, they mostly provide the subjective sense of certainty rather than an objective measurement in a field study. In this research, I utilize a field dataset from Yelp.com to test how different levels of certainty embedded in reviews influence persuasion, which is expected to strengthen the validity and generalizability of certainty effect.

When exploiting persuasion, source credibility has been of long-standing interest and inevitable attention (e.g., Tormala & Petty 2004). In particular, the source of a certain message can be deemed credible if he or she is either an expert who is perceived as professional with high ability or is a popular person who has a fame to ensure trustworthiness. Although the incremental effect of source credibility on the persuasiveness of a message is repeatedly studied (Forman et al, 2008; Otterbacher 2009; Baek et al. 2013), the further differentiation between source expertise effect and source popularity effect on persuasion of messages in degrees of certainty is not well examined. More importantly, expertise and popularity depict two different sides of reviewer features. Expertise highlights reviewer's rich experience and professional knowledge while popularity pinpoints his/her social networks in the form of fans following him/her. Chances are that people may perceive the certainty tone by an expert as persuasive while hesitating to accept the message of high certainty by a popular person. Thus, to clarify this question, I here examine two reviewer characteristics—i.e., expertise and popularity.

Moreover, customers often rely on categorization to identify and interpret the products or services provided by the organization. An organization can shape its identity by positioning itself into existing specific categories known as "niche width" (Kovács et al, 2013). Audiences tend to perceive an organization spanning multiple categories as with an ambiguous identity and a broad positioning while they perceive the organization concentrating on one specific category as with an explicit identity and a unique positioning—i.e., the authenticity (Hsu et al. 2009). Thus, organizational niche width can moderate users' ease and motivation to elaborate on reviews, thus having a role in affecting the perceived usefulness of reviews.

Existing research on review usefulness/helpfulness mostly focuses on its determinants, including review characteristics (e.g., emotions) (Yin et al, 2014) and reviewer characteristics (e.g., self-identity disclosure) (Forman et al, 2008). Although these studies contribute to the understanding of review helpfulness, firstly only limited information is known about the joint effects of review and reviewer features. In reality, these three—i.e., review content, reviewer and organizational information appear at the same time when customers read a review and try to evaluate its usefulness. In other words, the simultaneous existence suggests customers' interpretation of the three together. Specially, when exploiting persuasion, source credibility has been of long-standing interest and inevitable attention (e.g., Tormala & Petty 2004). Moreover, organizational characteristics are closely related with the content of reviews and viewers' interpretation of reviews. Owing to the actual situation of evaluating the usefulness of a review, I focus on the joint effects of the review, reviewer and organizational characteristics. Second, these studies mostly examine the role of product categorization (e.g., search vs. experience) (Mudambi & Schuff, 2010), and scant attention has been paid to examine the contextual factors from the perspective of an organization's positioning. Third, most of these studies operationalize review usefulness/helpfulness as a percentage of useful/helpful votes to the total number of votes. Only a few studies have directly modeled the count of usefulness/helpfulness votes that is of direct interest to social and behavioral scientists (Cao et al, 2011). And the method of predicting review usefulness in existing studies is yet to be well implemented. For example, to predict the count with a mean less than 10, certain studies (see, e.g., Racherla & Friske, 2012) adopted a linear regression (e.g., ordinary least squares regression) method that may create biased standard errors for significance tests (Gardner et al, 1995). Besides, some studies have failed to match the timeline between the review characteristics and the reviewer features (see, e.g., Wei et al, 2014). I intend to improve the estimation by using an appropriate method. Concretely speaking, I focus on the count of usefulness votes and estimate my model using zero-inflated negative binomial (ZINB) Poisson regression, predicting whether a review receives a useful vote and how many useful votes are obtained. Compared to the previous studies operationalizing review usefulness as a percentage, this prediction method can give a more comprehensive depiction of the accumulation of review usefulness votes. Well, it also matches with the dual-process theory which emphasizes the degree of persuasion induced via systematic or heuristic processing. In other words, the ZINB predictive method can not only capture whether users are persuaded—i.e., zero vote or not but also predict to what extent users are persuaded—i.e., number of usefulness votes.

Given the literature gaps identified in the previous paragraphs, my research aims to examine the interaction effects of review certainty, reviewer expertise and popularity, as well as organization niche width on review usefulness. First, I focus on the count of usefulness votes and estimate my model using zero-inflated negative binomial (ZINB) Poisson regression, predicting whether a review receives a useful vote and how many useful votes are obtained. Second, based on the existing literature, I add one more review textual feature, certainty, which is the subjective confidence or conviction of the expressed opinion about a product or service. This feature is measured by the frequency of certainty words occurring in a review, such as absolutely, must, always, and definitely. Third, I examine the interaction effects among the review certainty, reviewer characteristics and organization features. For reviewer features, I examine two social cues of a reviewer brought by the emergence of social media—reviewer expertise and popularity. For organization features, I consider the niche width of an organization based on the rationale that when an organization occupies wide demands, the recipients of the review have to validate intensive information. I integrate dual-process theory with social influence theory to explain my model.

Overall, I find that 1) reviewer popularity decreases the usefulness of the certainty-embedded review, and 2) the niche width of a restaurant magnifies the usefulness of the certainty-embedded review by the popular reviewer, whereas it mitigates the usefulness of the certainty-embedded review by the expert reviewer. My study provides important contributions to the e-WOM literature and extends dual-process theory. First, my study adds insights into the fast-growing stream of text mining studies that emphasize the role of textual characteristics in influencing consumer judgment (see, e.g., Kovács et al, 2013; Müller et al. 2016) by examining the certainty embedded in textual review content. Second, my findings supplement the review usefulness literature (see, e.g., Yin et al, 2014; Mudambi & Schuff, 2010; Forman et al, 2008) by verifying and identifying both the solitary and interaction effects of review certainty, reviewer popularity, expertise, and organization niche width. My study also adds to the current literature on review usefulness/helpfulness (see, e.g., Mudambi & Schuff, 2010) by directly examining the count of useful votes rather than the percentage of useful/helpful votes to the total number of votes using ZINB Poisson regression for estimation. Third, my findings contribute to the certainty literature (Clarkson & Tormala 2008; Rucker & Petty, 2004; Ryffel et al. 2014) by identifying three moderators to modify the persuasion effect of certainty from both reviewer and organization aspects. Fourth, my findings extend the dual-process theory by determining an additional peripheral cue, reviewer popularity, which has received less attention in dual-process research but is common in the context of social media. Finally, my findings that the joint effects of review certainty as well as reviewer popularity and expertise on review usefulness differ across organization niche widths contribute to the body of current research on organizational positioning (see e.g., Kovács et al., 2013; Dewan & Ramaprasad, 2012) and strengthen the understanding of how organizations may proceed concerning the generation of useful reviews. Practically, my findings offer actionable implications for managers of online review websites (e.g., Yelp.com) in critical information screening and selection, online retailers in providing guidelines for customer review-writing, and managers of social media platforms in enhancing the implementation of social media marketing.

1.2 Literature Review

Scholars have identified various determinants of review usefulness/helpfulness (see, e.g., Yin et al 2014; Mudambi & Schuff 2010; Forman et al 2008), including contextual attributes, reviewer characteristics, and review textual features. Existing studies mainly examine three key components of an online customer review: review, reviewer, and context.

One stream of studies examined numeric ratings and found that negative reviews are more useful in customer decision making than positive reviews (see, e.g., Chevalier & Mayzlin 2006). For example, Mudambi & Schuff (2010) identified that for experience goods, reviews with extreme ratings are usually perceived as less helpful than those with moderate ratings. However, the information embedded in reviews cannot be completely captured by numeric ratings (Resnick et al 2000). In recent years, scholars have directly investigated review text using text mining techniques. For instance, Korfiatis et al (2012) and Cao et al

(2011) identified the stylistic characteristics of a text (e.g., readability and word length) on the review helpfulness ratio. In terms of sentiment characteristics, Yin et al (2014) showed that the emotions embedded in a review influence customers' perception of review helpfulness and therefore proposed that anxiety-embedded reviews are more helpful than anger-embedded reviews. Moreover, Pan & Zhang (2011) adopted content analysis to capture the embedded innovativeness expressed in reviews and determined a curvilinear relationship between the expressed innovativeness and review helpfulness.

Aside from review textual factors, the characteristics of reviewers such as reviewer authorship (see, e.g., Forman et al 2008; Ghose & Ipeirotis 2011) and reviewer reputation (see, e.g., Otterbacher 2009) also influence review usefulness/helpfulness. Forman et al (2008) used the disclosure of self-identity information to explain the review helpfulness ratio and determined that such disclosure positively affects the review helpfulness ratio. Ghose & Ipeirotis (2011) demonstrated the positive effects of a reviewer's history, except for the authorship effects, on review helpfulness ratio. To investigate the review helpfulness on Amazon.com, Otterbacher (2009) viewed the reviewer's reputation in the community as one dimension of review quality and measured it with more than five metrics, such as the number of previous helpfulness votes received, number of total reviews written, and the "top reviewer" badge.

With regard to contextual factors, previous studies mostly focused on product characteristics. For example, Mudambi & Schuff (2010) focused on the moderating role of search versus experience goods in the impact of rating extremity on review helpfulness. Sen & Lerman (2007) examined the moderating role of utilitarian versus hedonic products in the relationship between review valence and review helpfulness, thereby suggesting that negative hedonic product reviews are less useful than those on utilitarian products.

Table 1 lists the relevant studies on review usefulness/helpfulness. Although these studies have been instrumental in enhancing my understanding of review usefulness, several interesting issues remain unanswered. First, the majority of the existing studies measure review helpfulness as the percentage of helpful votes to the total number of votes (see, e.g., Mudambi & Schuff 2010; Yin et al 2014), and only a few have directly modeled the count of helpfulness votes (e.g., Cao et al 2011). Considering that many online reviews in different websites have never obtained a single vote (Cao et al 2011), investigating the factors that predict zero-vote situations by focusing on the direct number of useful votes is important (Wei et al 2014). Second, most existing studies focus on solitary effects; hence, limited information is known about the joint effects of the three on review usefulness. To the best of my knowledge, only a few studies have directly investigated the role of review certainty as a predictor of review usefulness, except for Yin et al (2014), who indirectly tested the mediation mechanism of cognitive appraisal certainty to differentiate the effects of anger (high certainty) and anxiety (low certainty) on review helpfulness. Finally, contemporary studies emphasize the moderating role of product characteristics (i.e., search vs. experience), but few have looked into the matter from the perspective of the positioning of an organization (i.e., a generalist or a specialist). To fulfill the literature gap, this study determines the interaction impacts of review certainty, reviewer popularity and expertise as well as the niche width of an organization on review usefulness.

Table1. Summary of Relevant Studies on Review Usefulness/Helpfulness

Article	Data	Method	Usefulness/Helpfulness	Review Predictors	Reviewer Predictors	Context Predictors
Mudambi & Schuff (2010)	Amazon	Tobit regression	Percentage of helpful votes out of total votes	Length Rating extremity	N/A	Search vs. experience
Yin et al. (2014)	Yahoo!	Tobit regression	Percentage of helpful votes out of total votes	Anxiety Anger	N/A	N/A
Forman et al. (2008)	Amazon	Tobit regression	Percentage of helpful votes out of total votes	N/A	Self-identity disclosure	N/A
Ghose & Ipeirotis (2011)	Amazon	Tobit regression	Percentage of helpful votes out of total votes	Subjectivity Readability Spelling errors	Average helpfulness of reviewer' historical reviews	N/A
Korfiatis et al. (2012)	Amazon	Tobit regression	Percentage of helpful votes out of total votes	Length Readability	N/A	N/A
Otterbacher (2009)	Amazon	Simple linear regression	Percentage of helpful votes out of total votes	Topical relevancy Ease of understanding Believability Objectivity	Reviewer's reputation	N/A
Sen & Lerman (2007)	E-retailer website	OLS	Percentage of helpful votes out of total votes	Rating valence	N/A	Hedonic vs. utilitarian
(Pan & Zhang 2011)	Amazon	Logistic regression	Percentage of helpful votes out of total votes	Valence Age Length	Reviewer expressed innovativeness	Utilitarian vs. experience
Liu et al. (2008)	IMDB	Nonlinear regression	Percentage of helpful votes out of total votes	Writing style Timeliness	Reviewer expressed expertise	N/A
Baek et al. (2013)	Amazon	Hierarchical regression	Percentage of helpful votes out of total votes	Length Percentage of negative words to total word count	Rating inconsistency Reviewer ranking Reviewer real name	Search vs. experience High-priced vs. low-priced
Cao et al. (2011)	CNETD	Ordinal logistic regression	Number of helpfulness votes	Basic, stylistic and semantic characteristics	N/A	N/A
Racherla & Friske (2012)	Yelp	OLS	Number of usefulness votes	Elaborateness Valence	Self-identity disclosure Expertise Reputation	Search, experience vs. credence service
Wei et al. (2014)	Yelp	ZINB Poisson regression	Number of usefulness votes	Length Easy of understanding	Network centrality Elite badge	N/A
Ngo-Ye & Sinha (2014)	Yelp & Amazon	Hybrid text regression	Number of usefulness votes	Vector space model of review text	Reviewer engagement	N/A
Zhang et al.2010	Amazon	Binary logit model	Likelihood of helpfulness	Star rating	N/A	Promotion vs. prevention goal

1.3 Theoretical Foundation and Model

In this study, I select dual-process theory and social influence theory jointly as my theoretical foundation for the following reasons. First, dual-process theory is applied to illustrate the interaction between content and source of reviews, the shift between systematic processing and heuristic processing owing to the strength of the elaboration likelihood. By contrast, social influence theory goes deeply in the source effect and mainly explains two forms of social influence enforced by two characteristics of source—i.e., expertise and popularity. Given my research objective to examine the joint effect of review content and reviewer characteristics on persuasion, it is reasonable to consider both dual-process theory and social influence theory.

Second, these two theories build up my theoretical foundation together. On the one hand, the activation of informational influence is through internalization of information into one's own belief knowledge base and beliefs. And this internalization process needs highly involvement into systematic processing of message, which depends on elaboration likelihood. On the other hand, normative influence is enabled via identification process, which mostly leads to blindly following without effortful elaboration. And here I argue that the niche width of an organization which reviews are about may influence people's elaboration likelihood, because the niche width acts as a reflector of an organization's identity is used as reference.

1.3.1 Dual-process Theory

Dual-process theory hypothesizes that multiple factors influence the extent of how people think about the aspects of inter-communications, including the features of the source, recipient, and message (Petty & Cacioppo 1986). These theories indicate that when people are highly motivated and are able to assess a message, they tend to devote a huge amount of effort to elaborate the message by systematic processing so that the persuasion largely depends on message content, known as central route. When the elaboration decreases, peripheral cues such as source credibility become increasingly important in persuasion. This process is known as the peripheral route, also called heuristic processing. To the best of my knowledge, the empirical studies that apply dual-process theory mainly examine the factors that explain message quality or source credibility (Angst & Agarwal 2009). Source credibility is generally conceptualized as expertise and attractiveness. Source attractiveness has been mainly viewed as three interrelated aspects: familiarity, defined as the knowledge of the source by exposure (e.g., self-identity exposure) (Forman et al 2008); similarity, which refers to the supposed resemblance between the reviewer and recipient (e.g., friend vs. acquaintance) (Chu & Kim 20011); liking, known as the affection for the reviewer (e.g., physical beauty) (Maddux & Rogers 1980). In addition to reviewer expertise, I also examine the effects of reviewer popularity, an additional peripheral cue that has received less attention in dual-process research but is commonly found in the context of social media.

1.3.2 Social Influence Theory

Kelman (1961) indicated that social influence operates through three distinct processes, namely, compliance, identification, and internalization. Internalization occurs when people accept others' opinions and integrate them into their own belief systems. Compliance occurs when people publicly conform to others' opinions. Identification arises when people adopt others' opinions or behaviors to establish a relationship with the group. These processes can relate to two forms of social influence proposed by Burnkrant & Cousineau (1975), namely, informational and normative influences. Informational influence pertains to "the influence to accept information obtained from another as evidence about reality." Hence, such influence induces the acceptance of information corresponding to internalization process. Contrarily, normative influence is the tendency to conform to the expectations of others that can be attributed to either compliance or identification. For the identification process, an individual identifies himself/herself by adopting the behaviors or opinions that he/she perceives as representative of his/her reference groups. In this study, I use social influence theory to argue that reviewer's expertise influence goes through internalization process, whereas popularity influence via identification process.

Thus, normative influence highlights the identification mechanisms brought by the reviewer's social networks while informational influence emphasizes the internalization process activated by the signals of rich knowledge and experience. In this paper, expertise highlights reviewer's rich experience and professional knowledge while popularity pinpoints his/her social networks in the form of fans following him/her. Thus, I argue that reviewer's expertise influence goes through internalization process, whereas popularity influence via identification process.

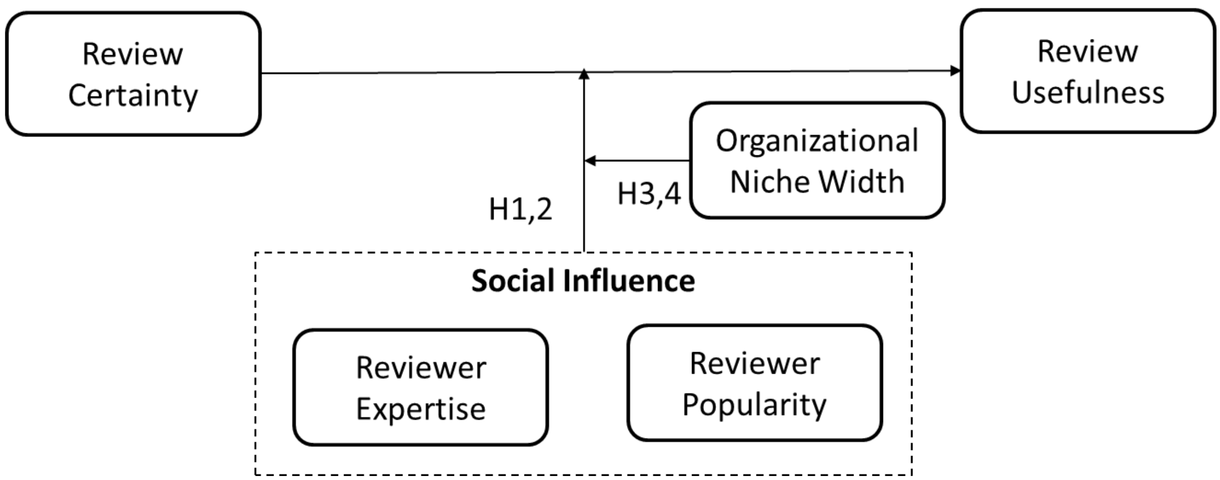
1.3.3 Research Model

To examine review usefulness, I integrate the dual-process theory with social influence theory to explain the underlying rationales behind my research model (Figure 1). As shown in Figure 1, I illustrate factors that predict review usefulness, namely, review certainty, reviewer expertise, reviewer popularity, and organization niche width. I argue that the effect of review certainty on review usefulness is moderated by the reviewer characteristics and that review certainty has different elaborative outcomes, depending on reviewer expertise and popularity.

I assert that a reviewer with a large number of followers is a popular reviewer while a reviewer with many reviews written is an expert. Reviewers with expertise are those who have rich experience in writing reviews on Yelp.com. Chances are that the reviewer who can write up high-quality reviews is more likely to capture users' attention and to be followed. Thus the reviewer popularity reflects the reviewer expertise to some extent.

In the current study, however, I treat popularity and expertise as two distinctive variables for the following reasons. First, popular reviewers may be well-known for other reasons than expertise, such as their writing 'funny' and 'cool' reviews. Thus their reviews may not be rich of information and valuable experience. Moreover, the large number of fans of a reviewer may attract more people to follow the focal reviewer, as a result, leading to herding effect. Thus, a reviewer with many fans is not necessarily the one with expertise. Second, on Yelp.com, people can directly detect Facebook friends and follow them. It is possible that the reviewer with much expertise may not be a social person with rich social networks. If so, he/she would not have as many Friends to transform from Facebook channel to Yelp channel, compared with the person who is possessed with strong social networks. Again, a reviewer with expertise is not necessarily the one with a lot of fans. Finally, and most importantly, popularity and expertise separately represents two kinds of social influence. The former depicts the normative social influence while the latter describes the informational influence. Normative influence highlights the identification mechanisms brought by the reviewer's social networks while informational influence emphasizes the internalization process activated by the signals of rich knowledge and experience. I thus treat popularity and expertise as two distinctive variables in my model.

I also consider the niche width of an organization and examine its moderation effects. Niche width is the construct to describe how many different categories an organization occupies and applicable to other contexts. As suggested by Kovács et al (2013), an organization can shape its identity by positioning itself into existing specific categories. An organization with a small niche width serves a small amount of the market, catering to specialized demands, which is often regarded as a specialist. By contrast, an organization with a big niche width serves a large amount of the market, catering to various demands, which is often regarded as a generalist. For example, Beauty Services which not only provide haircut service but also offer party make up service, can be called as a generalist. In contrast, if it only focuses on services related with hair, then it can be regarded as a hair specialist. Customers often rely on categorization to identify and interpret the products or services provided by the organization. Therefore, in my understanding, organizational niche widths can suggest the ease of elaboration on the review of the corresponding restaurant. These predictors and relationships are explained in detail in the subsequent section.



Note: * indicates the interaction effect between certainty and expertise and the interaction effect between certainty and popularity.

Figure1. Model of Review Usefulness

1.4 Hypotheses Development

1.4.1 Review Certainty

In online customer review websites (e.g., Yelp.com), customers can express their satisfactions (positive opinions) or dissatisfactions (negative opinions) toward a product or service. Regardless of whether they are satisfied or dissatisfied, the opinions of customers may vary in another dimension—in either confidence or ambiguity, which has been conceptualized as attitude certainty (Abelson 1988). Different from consumer behavior and social psychology research, in which attitude certainty is measured by asking a question—i.e., “How certain/convinced are you of your attitude?” (Rucker & Petty 2004), the present study directly captures the confidence of a reviewer’s opinion by calculating the frequency of certain words occurring in a review, such as must, absolutely, completely, and definitely.

There are mixed findings on the impacts of certainty on persuasion. Clarkson et al (2008) suggested that when controlling for valence and extremity, the opinions held with certainty are less conducive to systematic processing than those held with uncertainty. Yin et al (2014) demonstrated that the anger characterized by certainty induces reliance on peripheral cues, whereas the anxious reviewers appraised as uncertain tend to engage in effortful processing. Considering the point of these papers, certainty should negatively affect review usefulness for the lack of elaboration of certain messages. However, Sniezek & Van Swol (2001) specified that compared with advisors who express low certainty, those who hold high certainty are trusted more and their advice tend to be accepted. These mixed findings prove that certainty does not negatively influence persuasion consistently. In this study, I identify three moderators to modify the effect of certainty on review usefulness from both reviewer and organization aspects. I focus on reviewer expertise and popularity as reviewer characteristics, whereas I examine the niche width of an organization as one organization feature.

1.4.2 Review Certainty and Reviewer Expertise

Expertise is characterized as "an actor's ability to provide information to others because of his or her experience, education, or competence" (Biswas et al, 2006, p.19). Here, reviewer expertise is measured by the number of previous reviews written. Expert reviewers are those who are passionate in writing and post reviews on Yelp.com, thereby accumulating some knowledge in certain product categories. By contrast, non-expert reviewers post few reviews and lack experience in generating a review, let alone delivering professional knowledge to customers via evaluating products or services. I expect a positive impact of reviewer expertise on review usefulness. It is in nature that there is a positive direct impact of reviewer expertise on review usefulness. But in this study, I focus on the moderating role of reviewer expertise on effect of review certainty on review usefulness. Thus, the logic is that the expertise of a review could facilitate viewers to be highly involved in elaboration of the certain comments and thus more likely to be persuaded by the certainty-embedded review. And the underlying assumption is that the more cognitive effort involved in understanding the certain points in the review, the more information will be internalized as individual knowledge and belief, and finally the more persuasive the review will be thought of, thus more useful. Here my interest is in the moderating role of reviewer expertise on effect of review certainty on review usefulness.

When the reviewer is an expert who has written many reviews, social influence occurs through internalization. Internalization transpires when the opinions presented by an expert seem useful for the solution of a particular problem. The rich experience of reviewers in writing reviews signals their cognitive effort expended in generating the certainty-embedded review. In this case, the recipients of a review believe that a reviewer with rich review-writing experience and relative knowledge is willing to spend effort to provide evaluations and ensure recommendation certainty. In other words, reviewer expertise increases the recipients' perception of a reviewer's cognitive efforts; such an increase fosters the recipients' increased involvement and elaboration, thereby driving the usefulness of the certainty-embedded review.

H1: Reviewer expertise positively moderates the effect of review certainty on the usefulness of the review such that the higher the expertise, the greater impact of certainty on review usefulness.

1.4.3 Review Certainty and Reviewer Popularity

In marketing, a celebrity endorser is defined as "any individual who enjoys public recognition and who uses this recognition on behalf of a consumer good by appearing with it in an advertisement" (Biswas et al, 2006, p.18). In this study, I examine celebrity endorsements by focusing on popular reviewers on online review websites (e.g., Yelp.com). Popular reviewers are those who are followed by a large number of fans, whereas the unpopular ones are followed by only a few fans. I expect a positive impact of reviewer popularity on review usefulness. It is not surprising that there is a positive effect of reviewer popularity on review usefulness. But in this study, I hypothesize a negative moderating the role of reviewer popularity on the effect of certainty on review usefulness. That is because a recipient of a review posted by a popular reviewer who is surrounded by hundreds of fans is supposed to engage in mindless, heuristic processing that requires minimal direct thought of message, leading to less deliberation of the certain points or comments in the review. As a result, recipients are less likely to be persuaded by the certainty-embedded review and perceive it as useful. Here I focus on the moderating role of reviewer popularity on the impact of certainty on review usefulness.

When an individual is extremely popular online, social influence occurs via identification. In particular, this process takes place when a peer attempts to establish or maintain the identity associated with the popular person. Applying identification influence to my context, a recipient of a review posted by a popular reviewer who is surrounded by hundreds of fans is supposed to engage in mindless, heuristic processing that requires minimal direct thought of message. In other words, when recipients encounter a review posted by a popular reviewer with many fans, they tend to rely on superficial cues rather than conduct a careful deliberation and elaboration, thereby leading to less usefulness of the certainty-embedded review.

H2: Reviewer popularity negatively moderates the effect of review certainty on the usefulness of the review such that the higher the popularity, the weaker impact of certainty on review usefulness.

1.4.4 Moderating Role of Organization Niche Width

In a nutshell, the elaboration of a review on a generalist which covers multiple categories is more demanding and effortful than that of a review of a specialist concentrating on one category. This is because customers often rely on categorization to identify and interpret the products or services provided by the organization. A business can shape its identity by positioning itself into existing specific categories known as "niche width" (Kovács et al, 2013). For example, restaurants spanning multiple cuisines (e.g., serving both Chinese dishes and Japan sushi) often display features richer than those of the single-cuisine restaurants (e.g., only serving Hamburger) (Kovács et al, 2013). Audiences tend to perceive an organization spanning multiple categories as with an ambiguous identity and a broad positioning while they perceive the organization concentrating on one specific category as with an explicit identity and a unique positioning—i.e., the authenticity (Hsu et al. 2009). Thus, organizational niche width can moderate users' ease and motivation to elaborate on reviews, thus having a role in affecting the perceived usefulness of reviews.

Following the above elaboration on the notion of organization niche, I argue that it is cognitively more demanding for a user to evaluate a generalist than a specialist because of the ambiguous identity and the broad positioning. This will in turn lead to the user's less motivation of the elaboration and evaluation of the review of a generalist. That is, the more diversified categories an organization have, the less motivation for users to absorb and digest the information from the reviews about this organization. This would inhibit the internalization of the reviews into self-belief or persuasion.

As discussed in H1, reviewer expertise enhances the effect of certainty comments on review usefulness via informational influence by the internalization process. The relevance and ease of understanding of the information provided will motivate people to devote cognitive efforts to systematic processing and knowledge absorption—i.e., the internalization of the strong opinion. However, reviews about a broad-positioning organization—i.e., greater organization niche width—will weaken the enhancing effect of reviewer expertise due to the difficulty of elaboration on the information. Thus, the moderating effect of expertise will be weakened by organization niche width.

There are two reasons for the above argument. First, the scope of a review done by a reviewer may reflect the scope of the categories an organization covers. After all, the more categories included, the more category-associated products can be commented or recommended in the review. The complexity of a review of a generalist prevents the internalization. Thus, I argue that when an expert reviewer recommends a narrow-positioning organization, the recipients can easily understand a certain opinion and transform the information obtained from the review into their own knowledge, thereby increasing their level of understanding. Second, even if a review may not necessarily cover all categories of a generalist, a user may still be less motivated to internalize the review. For instance, the category(s) covered in a expert's review may not be the one of a reader's interests. Or, it is difficult to elaborate on the broad information provided by the expert to verify authenticity (Hsu et al. 2009).

Therefore, the positive moderation effect of expertise in the two-way interaction will be conversely negatively moderated—i.e., a negative three-way interaction effect—when organization niche width plays its role in influencing readers' judgement of review usefulness.

H3: There is a negative interaction impact of organization niche width, review certainty and reviewer expertise on usefulness of the review such that the certainty-embedded review written by an expert reviewer is more useful for an organization operating within fewer categories.

In contrast, reviewers with high popularity have normative influence by the identification process that does demand effortful elaboration on the review content. Popularity weakens a reader's motivation to elaborate on a review to verify the certainty comments. However, the broad range of categories of a generalist and the vagueness of its authenticity will prevent readers' simply following the popular reviewer to verify the certainty comments. Instead, systematic process will be likely employed because readers are motivated to carefully elaborate on the certainty comments in order to figure out the specific category of products commented. In other words, for reviews about a broad-positioning organization (e.g., a restaurant that serves multiple cuisines, including Chinese hot pot, Japanese sushi, and American sandwiches), recipients tend to be more cognitively involved in understanding the content of reviews to first determine the category of product(s) with which reviewers are satisfied or dissatisfied definitely and then internalize the digested information as evaluation reference, instead of blindly following popular reviewers without effortful thinking. By contrast, for reviews of a narrow-positioning organization—e.g., a restaurant that only serves Chinese hot pot, the ease of evaluating and understanding makes recipients completely reliant on popular reviewers. As a result, for narrow-positioning organizations, recipients simply follow certainty-embedded recommendations without careful reading and deep thinking of the review content.

Thus, the negative moderation effect of popularity in the two-way interaction will be further negatively moderated—i.e., a positive three-way interaction effect—when organization niche width plays its role in influencing readers' judgement of review usefulness.

H4: There is a positive interaction impact of organization niche width, review certainty and reviewer popularity on usefulness of the review such that the certainty-embedded review written by a popular reviewer is more useful for an organization operating within more categories.

1.5 Data

1.5.1 Data Collection

The research context is Yelp.com, a popular online review website founded on October 2004. Yelp covers a broad range of 22 product and service categories, such as restaurants, shopping, beauty, and spas, each of which contains subcategories. For example, the "restaurants" category includes 75 subcategories, such as Chinese, Japanese, Pizza, and Sandwiches. Some restaurants are titled with a single subcategory, whereas others occupy multiple subcategories. This categorization is accomplished by the website, sometimes with the consultation with restaurants. In addition, the Yelp interface provides information about the reviewers to help provide quick evaluations about the reviewer and his/her reviews. Displayed below each reviewer name and registered city are the number of reviews and number of friends that reviewer has; these elements indicate how heavily the reviewer is involved in the website. A reviewer becomes someone's fan or friend if his/her review is appreciated by that person. The key difference between friends and fans is that the former is visible to see their profile, whereas the latter is anonymous. On this website, anyone (with or without an account) can read a written and published review and give a vote, including "useful," "funny," or "cool."

I adopt the Yelp Academic Data Set (https://www.yelp.com/academic dataset) released on January 2014 to establish my own research sample using the SAS software. First, I select restaurants as the target research object because a restaurant is a typical experience good, the quality of which cannot be thoroughly inspected before purchase. Second, I select restaurants with active listings on Yelp on January 2014; Yelp only includes restaurants that remain operational at the time of searching and viewing. Third, I only focus on reviews within the three months before the released time, that is, from Oct 2013 to Dec 2013[2]. Conceptually, I choose the short 3-month period as my research time window based on the following justifications. Reviewer characteristics—i.e., here point at the number of reviews written and the number of fans following the reviewer, change over time. That is not only the usefulness votes on the review go up but also the reviewer may write up more reviews and attract more fans since the review is posted out on Yelp.com. And more importantly, there exists potentially temporal asymmetry between review content and reviewer characteristics. Given that my focal variables—i.e., number of reviews written and number of fans, are observed statistically at the time of data collection—i.e., Jan., 2014. Thus, to match the timeline between the content and source characteristics, I focus on this short observation time window and assume no significant changes in reviewer characteristics. Finally, I only examine reviews comprising more than 50 words because for texts with less than 50 words, the content analysis obtained from the linguistic inquiry word count (LIWC) program is of low credibility (http://liwc.wpengine.com/how-it-works/). Thus, the entire sample includes 10,097 reviews written by 6,191 reviewers of 2,383 restaurants in the state of Arizona, U.S.A. from Oct 2013 to Dec 2013.

1.5.2 Data Preparation

To capture the textual characteristics of the reviews, Iconduct content analysis using the Linguistic Inquiry and Word Count (LIWC) software program (Pennebaker et al, 2007), the reliability and validity of which have been extensively investigated (Pennebaker et al, 2007).

[2] Empirically, I also test my hypotheses using the dataset 6 months before data collection. The results are shown in the Appendix A, which partially provided significant evidence for the research model. The results gave support for the rationality in the time window choice. Considering the potential asymmetry between review content and reviewer characteristics, the current time window works well.

This technique has been recently and frequently used in IS and marketing research. For example, with the use of LIWC, Humphreys (2010) tracked the creation of casino gambling markets by analyzing newspapers and Yin et al. (2014) captured the specific emotions "anger" and "anxiety" embedded in reviews.

The LIWC application relies on an internal default dictionary that defines which words should be counted in the target text files. The dictionary is composed of almost 4,500 words and word terms, each of which defines one or more word processes and categories. There are 26 word categories representing linguistic processes with general descriptor categories (e.g., total word count and words per sentence) and standard linguistic dimensions (e.g., percentage of words in the text that are pronouns, articles, and auxiliary verbs); 32 word categories tapping psychological processes that include social processes (e.g., family, friends, and humans), affective processes (positive emotion, negative emotion, anger, anxiety, and sadness), and cognitive processes (e.g., certainty, insight, and causation); 7 personal concern categories (e.g., work, home, and leisure activities); 3 spoken categories (e.g., assents, fillers, and nonfluencies); and 12 punctuation categories (e.g., periods and commas). For the processing of each word, LIWC searches its dictionary for a match, and if a match occurs, the corresponding category scale for that word would be incremented. At the end of this process, a final score is added to each category, representing the percentage of associated words in the text sample matching that category. For example, the word "annoyed" would be assigned to six word categories: anger, negative emotion, overall affect, ad, verb, and past tense verb. Hence, if this word is found in the target text, each of these six categories scale scores will be incremented. In addition, LIWC dictionary allows for any target word that matches the first five letters of the word to be counted as an ingestion word.

In this study, I conduct content analysis on the text of each review entered for each restaurant using LIWC, which calculates the total frequency of the dictionary words that appear in a category divided by the total number of words in the review, to determine the percentage of a review that falls into different categories. Here I mainly focus on affective processes (e.g., positive emotion, negative emotion, anger, anxiety), cognitive processes (e.g., certainty) and linguistic processes (e.g., word count). The total number of words in a review is measured as the word length of the review. Given that the ease of reading is important for review elaboration, I also calculate the readability of each review by using the Gunning fog index. The word category "certainty" represents the key variable of interest. The certainty category includes 83 associated words, such as absolute, certain, clearly, commit, completely, confidence, fact, must, definitely, totally, and every. Therefore, certainty is measured as the percentage of the number of certainty-related words divided by the total number of words in a review. In addition, I also calculate positivity, negativity, anger, and anxiety by using the number of positive emotion-, negative emotion-, anger-, or anxiety-related words to divide the total number of words.

1.5.3 Variables and Measurement

Dependent Variable. Below each review, Yelp.com presents the question "Was this review…?" along with "useful," "funny," and "cool" options. In this study, I only focus on the "useful" option. A review that has received at least one useful vote displays the number of useful votes immediately beside the "useful" icon. Thus, review usefulness is measured as the number of useful votes; a high value of usefulness indicates a useful review. Review usefulness is my dependent variable of interest.

Independent Variables. I regard review certainty, reviewer expertise, reviewer popularity, restaurant niche width, and their interactions as predictors of review usefulness. The certainty embedded in a review is measured with LIWC as explained in "Data Preparation" section. Table 2 lists the measures of the other three variables.

Control Variables. The control variables are the other characteristics of review, reviewer, and restaurant. For review characteristics, I control for readability (Ghose & Ipeirotis 2011), review length, review rating and squared terms of rating (Mudambi & Schuff 2010), anger, anxiety (Yin et al 2014), and review timespan (Racherla & Friske 2012). To accurately control for the emotional valence of textual reviews, I also control for the percentage of the positive and negative emotional words in a review calculated by LIWC. For reviewer characteristics, I control for status, time spent on Yelp, and average rating (Ghose & Ipeirotis 2011). Finally, for restaurant characteristics, I control for the price range of a restaurant and the age of a restaurant until the data collection time. Given that the establishing date of a restaurant is unavailable for use, I view the date of the first review of a restaurant as a proxy for the opening date of the restaurant. In addition, I also control for the average rating (i.e., reputation of the restaurant) and the total number of reviews (i.e., popularity of the restaurant) a restaurant has received. Tables 2, 3, and 4 present the measures, descriptive statistics, and correlation matrix of all variables, respectively.

Table2. Measurement of Variables

Variable Type	Variable Name	Measures
Dependent Variables	Review usefulness	Number of usefulness votes
Independent Variables	Review certainty	(Certainty-related words/total words in a review) *100
	Reviewer expertise	Number of previous reviews written by a reviewer
	Reviewer popularity	Number of fans of a reviewer
	Restaurant niche width	Number of cuisines a restaurant occupies
Control Variables	Rating	Star rating (1-5) of a review
	Length	Number of words in a review
	Readability	Gunning Fog Index=0.4*(average words per sentence+ count of hard word for each 100 words), where a "hard" word here is defined as a word with more than six characters. Note that the larger readability, the harder to read the review.
	Anger	(Anger-related words/total words in a review) *100
	Anxiety	(Anxiety-related words/total words in a review) *100
	Positivity	(Positive emotion-related words/total words in a review) *100
	Negativity	(Negative emotion-related words/total words in a review) *100
	Timespan (weeks)	Number of weeks elapsed since a review posted
	Reviewer status	A dummy variable, titled as "elite" or not
	Reviewer yelping time (weeks)	Number of weeks elapsed since a reviewer registered
	Reviewer average rating	Average star rating of previous reviews written by a reviewer
	Restaurant price	Price level ranging from $, $$, $$$ to $$$$
	Restaurant reputation	Average star rating of a restaurant
	Restaurant popularity	Total number of reviews obtained by a restaurant
	Restaurant age(weeks)	Date of data collection minus date of the first review of a restaurant

Table3. Descriptive Statistics

Variables	Mean	Std. Dev.	Min	Max
Review usefulness	0.71	1.64	0.00	48.00
Review certainty	1.60	1.32	0.00	11.43
Reviewer expertise	71.30	162.44	1.00	2110.00
Reviewer popularity	3.30	17.74	0.00	569.00
Restaurant niche width	1.80	0.83	1.00	3.00
Rating	3.76	1.31	1.00	5.00
Length	154.68	101.40	50.00	530.00
Anger	0.25	0.54	0.00	10.94
Anxiety	0.14	0.40	0.00	6.06
Positivity	5.48	2.79	0.00	20.37
Negativity	1.04	1.18	0.00	12.73
Readability	12.71	5.04	3.44	127.74
Timespan	10.78	3.77	4.35	17.32
Reviewer status	0.15	0.35	0.00	1.00
Reviewer yelping time	119.27	84.90	4.35	473.63
Reviewer average rating	3.75	0.79	1.00	5.00
Restaurant reputation	3.79	0.52	1.00	5.00
Restaurant popularity	144.52	165.89	3.00	1124.00
Restaurant age	247.01	134.86	12.12	473.57

Table 4. Correlation of Variables

Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
1. Usefulness	**1**																		
2. Certainty	-0.01	**1**																	
3. Expertise	0.33***	0.00	**1**																
4. Popularity	0.35***	-0.00	0.63***	**1**															
5. Niche width	-0.01	0.01	0.02	0.01	**1**														
6. Status	0.32***	0.01	0.52***	0.33***	0.01	**1**													
7. Rating	-0.03**	0.03**	-0.00	0.02	0.02*	0.02*	**1**												
8. Length	0.02	-0.09***	0.01	0.01	-0.00	0.01	-0.00	**1**											
9. Anger	0.01	0.01	-0.01	-0.00	-0.01	0.02*	-0.04***	0.03***	**1**										
10. Anxiety	0.01	-0.00	0.00	0.01	-0.01	0.00	-0.04***	0.02	0.05***	**1**									
11. Positivity	0.00	0.18***	0.01	0.00	0.04***	0.01	0.10***	-0.28***	-0.13***	-0.10***	**1**								

12. Negativity	0.01	0.01	-0.01	-0.00	-0.02	0.00	-0.08***	0.02*	0.55***	0.42***	-0.24***	**1**							
13. Readability	0.00	0.00	-0.00	-0.00	0.01	-0.01	0.01	0.11***	0.02	0.01	-0.05***	0.02	**1**						
14. Timespan	-0.00	0.02	0.02	0.01	0.01	-0.02	0.02	-0.01	0.00	0.00	-0.01	0.00	-0.00	**1**					
15. Average rating	0.00	0.00	0.00	0.02*	0.01	0.05***	0.57***	0.01	-0.02*	-0.03***	0.05***	-0.05***	0.01	0.03*	**1**				
16. Yelping time	0.15***	0.01	0.40***	0.24***	0.01	0.36***	-0.01	0.01	-0.01	0.00	0.01	-0.02*	-0.01	-0.01	-0.02	**1**			
17. Restaurant reputation	0.04***	0.07***	-0.01	0.00	0.10***	0.01	0.39***	0.01	-0.11***	-0.11***	0.19***	-0.19***	-0.01	0.01	0.22***	0.05***	**1**		
18. Restaurant popularity	-0.01	0.03**	-0.01	0.00	0.18***	0.00	0.11***	0.07***	-0.03**	-0.03***	0.11***	-0.06***	0.02*	0.02*	0.06***	0.04***	0.28***	**1**	
19. Restaurant age	-0.07***	0.04***	-0.02	-0.02	-0.01	-0.02*	-0.01	0.01	0.02*	-0.04***	-0.02	-0.00	0.00	0.03**	-0.02*	-0.03**	-0.04***	0.35***	**1**

$^{***}p<0.001$; $^{**}p<0.01$; $^{*}p<0.05$

Note: Regarding the high correlations between some variables, I use VIF (variance inflation factor) to measure the severity of multi-collinearity and find that the VIF values of all the variables are less than 5. Hence, the multi-collinearity issue will not create a problem for my estimation results.

1.6 Estimation and Results

1.6.1 Model Specification

When modeling count data—i.e., the number of useful votes of a review, I have to consider two issues: first, whether or not a useful vote is given, and second, how many useful votes are made on the condition of one useful vote given. Thus, in this study, I both consider the likelihood of zero useful votes made and the number of useful votes given on the condition of one useful vote obtained. I propose the use of ZINB model for estimation primarily because it is appropriate to model the over-dispersed count data with excess zeros (Coxe et al, 2009). My data not only demonstrate over-dispersion with the variance of count of "useful" votes larger than the mean (as depicted in Table 3, variance (usefulness) =2.69, mean (usefulness) =0.71) but also contain excess zeros with a large percentage of reviews (63.22%) obtaining no useful vote. With the use of the ZINB Poisson model, I can test two-stage models in which the logit and the standard negative binomial (NB) model are estimated jointly; the former estimates the probability of a review to receive zero useful votes and the latter predicts the conditional number of useful votes.

The logit model that determines whether a review gains a useful vote is specified in Equation (1), including restaurant features (i.e., restaurant niche width, price, reputation, popularity, and restaurant age), reviewer characteristics (i.e., reviewer yelping time, expertise, popularity and status), and review valence (i.e., rating and its squared terms). In line with the existing empirical studies, customers' perceptions of online reviews (e.g., helpful or not) are largely determined by review ratings (e.g., rating valence and extremity) (Mudambi & Schuff 2010), reviewer traits (e.g., number of past reviews written) (Ghose & Ipeirotis 2011), reviewer deemed as "elite" or not, number of friends obtained (Racherla & Friske 2012), and time the reviewer has spent on Yelp (Wei et al 2014), and organization features (i.e., store reputation and popularity) (Yin et al 2014), restaurant price, and niche width (Kovács et al 2013). The interaction terms are excluded from the logit model because I believe that the interactions among review, reviewer, and contextual features are more relevant to how many useful votes are given than the probability of not giving any useful votes.

On the other hand, the NB model presents the number of usefulness votes as specified in Equation (2). The model estimates the effects of review certainty, reviewer expertise, reviewer popularity, restaurant niche width, and interactions among them as well as the control variables in line with the existing literature. The ZINB model allows for the existence of different sets of predictors in the logit and the standard NB models (Coxe et al 2009).

$$Logit(Y_{i,j,k}{}^{*}) = \alpha_0 + \alpha_1 Rest_j + \alpha_2 Reviewer_k + \alpha_3 \mathrm{Rating}_{i,j,k} + \alpha_4 \mathrm{Rating}_{i,j,k}{}^{2} + \varepsilon_{i,j,k}{}' \quad (1)$$

$$Log(Y_{i,j,k}) = \beta_0 + \beta_1 Cer_{i,j,k} + \beta_2 Rev_k + \beta_3 Fans_k + \beta_4 Nichew_j + \beta_5 TwoInter_{i,j,k} + \beta_6 ThreeInter_{i,j,k} + \beta_7 Contr_{i,j,k} + \varepsilon_{i,j,k} \quad (2)$$

where

$\mathrm{Y}_{i,j,k}{}^{*}$ is the probability of zero usefulness votes on review i of business j by reviewer k;

$Y_{i,j,k}$ is the expected number of usefulness votes on review i of business j by reviewer k;

Rest_j is a matrix of variables about restaurant j, including $\mathrm{restaurant\_reputaion}_j$, $\mathrm{restaurantage}_j$, $\mathrm{restaurant\_popularity}_j$, *and* price_j;

$Reviewer_k$ is a matrix of variables about reviewer k, including status_k, $\mathrm{Yelping\_time}_k$;

$\mathrm{Cer}_{i,j,k}$ is the certainty of review i of business j by reviewer k;

Rev_k is the number of reviews written by reviewer k;

$Fans_k$ is the number of fans of reviewer k;

$NicheW_j$ is the niche width of restaurant j;

$\mathrm{TwoInter}_{i,j,k}$ is a matrix of two-way interaction terms, including $\mathrm{Cer}_{i,j,k}*\mathrm{Fans}_k$, $\mathrm{Cer}_{i,j,k}*\mathrm{Rev}_k$, $\mathrm{Fans}_k * \mathrm{Rev}_k$, $\mathrm{Fans}_k*\mathrm{NicheW}_j$, $\mathrm{Rev}_k*\mathrm{NicheW}_j$, $\mathrm{Cer}_{i,j,k}*\mathrm{NicheW}_j$;

$\mathrm{ThreeInter}_{i,j,k}$ represents a matrix of three-way interaction terms, including $\mathrm{Cer}_{i,j,k}*\mathrm{Fans}_k*\mathrm{NicheW}_j$, $\mathrm{Cer}_{i,j,k}*\mathrm{Rev}_k*\mathrm{NicheW}_j$;

$\mathrm{Controls}_{i,j,k}$ represents a matrix of control variables for review i of business j by reviewer k , including $\mathrm{rating}_{i,j,k}$, squared term of $\mathrm{rating}_{i,j,k}$, $\mathrm{length}_{i,j,k}$, $\mathrm{readability}_{i,j,k}$, $\mathrm{anger}_{i,j,k}$, $\mathrm{anxiety}_{i,j,k}$, $\mathrm{positivity}_{i,j,k}$, $\mathrm{negativity}_{i,j,k}$, $\log(\mathrm{timespan}_{i,j,k})$, $\mathrm{Yelping\_time}_k$, status_k ; $\mathrm{Avguser\_rating}_k$, $\mathrm{restaurant\_reputaion}_j$, $\mathrm{restaurantage}_j$, $\mathrm{restaurant\_popularity}_j$, price_j;

$\varepsilon_{i,j,k}{}'$ is the residual error term of equation (1).

$\varepsilon_{i,j,k}$ is the residual error term of equation (2);

Note that Rev_k and Fans_k are treated as static variables—i.e., I assume no significant changes in these reviewer characteristics in my three-month short observation time window.

1.6.2 Estimation Results

My ZINB model is estimated in SAS using Proc Genmod. I control for the variable length of the observation periods of the reviews by including the natural log of review timespan using the offset option; then, its regression coefficient is equal to 1 (Allison 2012). Vuong test is conducted to compare the ZINB model with the standard NB model, and the results for Schwarz Adjusted Statistic are obtained ($z=8.2199$, $p<.001$). The findings suggest that the ZINB model is a significant improvement over the standard NB model. The scaled Pearson chi-square statistic is significantly different from 1 (scaled Pearson $X2=10658.0673$, $p<.001$), thereby providing evidence for over-dispersion.

To clarify my hypothesized effects, I use stepwise regression with four blocks of variables: controls (Model 1), linear effects (Model 2), two-way interaction effects (Model 3), and three-way interaction effects (Model 4), each of which estimates the logit and NB models jointly. Tables 5 and 6 demonstrate the estimation results.

The same set of variables is used for logit regression across all four models. As shown in Table 5, regardless of the models used, the logit results show that restaurant reputation constantly decreases the probability of a review obtaining zero useful votes. In other words, the reviews about restaurants with high star ratings tend to receive useful votes. The reason is that a restaurant with a high rating (e.g., five stars) is perceived by customers of high quality and easily attracts considerable attention, and its reviews tend to be read and evaluated. In addition, the results of the logit model show that the reviewer's time spent on Yelp also decreases the probability of a review obtaining zero useful votes. In other words, the probability of a review to receive zero useful votes is low when the reviewers have been associated with Yelp.com for a long period of time (starting with their registration on Yelp). This result is attributed to the logic that the longer the time that reviewers stay on a website, the more familiar they become with the environment and the regulations of the website. Such familiarity helps the reviewers easily capture the preferences of customers on the website and write useful reviews.

Table 5. ZINB Estimation Results (Logit Model)

Variables	Logit Model Estimating Zero Useful Votes			
	Model 1	Model 2	Model 3	Model 4
Rating	4.3360	1.7717	1.6481	1.3800
Squared rating	-0.4492	-0.1565	-0.1412	-0.1116
Reviewer Status	-3.4540	-19.3357	-19.3483	-20.4475
Reviewer Yelping Time	-0.0477**	-0.0365***	-0.0356***	-0.0265***
Restaurant Popularity	0.0021*	0.0014	0.0013	0.0012
Restaurant Reputation	-1.3536***	-0.9675***	-0.9605***	-0.9190***
Restaurant age	0.0008	0.0012	0.0012	0.0012
Price: $	-1.8266	-1.4172	-1.4925	-1.7238*
Price: $$	-1.9027	-1.5013	-1.5838	-1.7720*
Price: $$$	-2.1551	-1.8342	-1.9244	-2.0540*
Price: $$$$	0.0000	0.0000	0.0000	0.0000

$^{***}p<0.001$; $^{**}p<0.01$; $^{*}p<0.05$

For negative binomial regression, as shown in Table 6, I first estimate Model 1. The control effects are partially consistent with the prior literature. First, the negative reviews (β = –0.4192, p<.001), especially with extreme ratings (β = 0.0478, p<.001), are more useful than the positive ones. This result is consistent with the findings obtained by Mudambi & Schuff (2010). Second, a review about a restaurant with great reputation (β = 0.2098, p<.001) and that is opened only recently (β = –0.0012, p<.001) is considered more useful than those with low reputation and have long been operating. Third, the coefficient of reviewer status (β = 1.3686, p<.001) is significant and positive, thereby suggesting that the reviews posted by the “elite” are predicted useful; such a finding is in line with the observations of Ghose & Ipeirotis (2011).

Model 2 validates the linear effects of review and reviewer characteristics on review usefulness. The coefficients of reviewer popularity (β = 0.0162, p<.001) and expertise (β = 0.0006, p<.01) are significant and positive, thereby suggesting that the reviews posted by well-known reviewers or those with rich posting experience are predicted to be useful. Furthermore, the coefficient of restaurant niche width (β = –0.0650, p<.01) is significant and negative, thereby indicating that the reviews about narrow-positioning restaurants are more useful than those on broad-positioning restaurants.

Model 3 tests the hypotheses H1 and H2. The coefficient of the two-way interaction between review certainty and reviewer popularity ($\beta = -0.0040$, $p<.01$) is significant and negative, which suggests that H2 is supported. However, I fail to secure significant evidence for H1 that depicts the interaction between review certainty and reviewer expertise ($\beta = 0.0001$, $p=0.1940$).

Model 4 tests the hypotheses H3 and H4. The coefficient of the three-way interaction among review certainty, reviewer expertise, and restaurant niche width ($\beta = -0.0004$, $p<.01$) is significant and negative, whereas the coefficient of the three-way interaction among review certainty, reviewer popularity, and restaurant niche width ($\beta = 0.0061$, $p<.001$) is significant and positive, thereby supporting both hypotheses.

Table 6. ZINB Estimation Results (NB Model)

Variables	NB Model Estimating No. of Useful Votes			
	Model 1	Model 2	Model 3	Model 4
Model 1: Control effects				
Rating	-0.4192***	-0.5151***	-0.5219***	-0.5586***
Squared rating	0.0478***	0.0624***	0.0635***	0.0712***
Length	0.0001	0.0002	0.0002	0.0002
Anger	-0.0008	0.0265	0.0282	0.0277
Anxiety	-0.0072	0.0197	0.0168	0.0117
Positivity	-0.0022	-0.0033	-0.0034	-0.0041
Negativity	0.0248	0.0085	0.0096	0.0046
Readability	0.0025	0.0036	0.0035	0.0036
Log(Review Timespan)#	1	1	1	1
Reviewer Status	1.3686***	0.9641***	0.9642***	0.5732***
Reviewer Yelping Time	0.0006*	-0.0009**	-0.0009**	-0.0014***
Reviewer Average Rating	0.0447	0.0683*	0.0695*	0.0701*
Restaurant Reputation	0.2098***	0.2274***	0.2233***	0.1942***
Restaurant Popularity	0.0000	0.0001	0.0001	0.0001
Restaurant age	-0.0012***	-0.0011***	-0.0011***	-0.0011***
Price: $	-0.2994	-0.0962	-0.1265	-0.3103
Price: $$	-0.1108	0.0896	0.0553	-0.1273
Price: $$$	-0.0786	0.0817	0.0463	-0.1058
Price: $$$$	0.0000	0.0000	0.0000	0.0000
Model 2: Linear effects				
Review Certainty		0.0075	0.0106	0.0124
Reviewer Expertise		0.0006**	0.0005**	0.0012***
Reviewer Popularity		0.0162***	0.0156***	0.0423***
Restaurant Niche Width		-0.0650**	-0.0639**	-0.0658**
Model 3: Two-way Interaction effects				
Certainty*Expertise (**H1**)			0.0001	-0.0001
Certainty*Popularity (**H2**)			-0.0040**	0.0014
Certainty*Niche width				0.0174
Expertise*Popularity				-0.0000***
Expertise* Niche width				-0.0000
Popularity* Niche width				0.0010
Model 4: Three-way Interaction effects				
Certainty*Expertise*Niche width (**H3**)				-0.0004**
Certainty*Popularity*Niche width (**H4**)				0.0061***

***p<0.001; **p<0.01; *p<0.05

#As suggested by Allison (1999), I included the natural log of the review timespan as a predictor with regression coefficient equal to 1 with the purpose of incorporating variable observation periods while maintaining the Poisson error structure of the data.

1.7 Discussion

1.7.1 Summary of Results

I examine the joint effects of review, reviewer, and organization characteristics on review usefulness by building on dual-process theory. Through the application of the paradigm of expert and celebrity endorsements (Biswas et al 2006) explained by social influence theory, I offer a conceptualization of what constitutes a useful review. Obtained by utilizing restaurant reviews from Yelp.com with ZINB regression, my empirical results provide support for the model and most of my hypotheses. The certainty-embedded review receives fewer usefulness votes when written by a popular reviewer followed by many fans than when written by a less popular reviewer. I conjecture that this condition is attributed to the fact that the signal of popularity (number of fans) tends to inspire customers to conduct mindless heuristic processing, thereby mitigating their cognitive efforts to understand the review content. For the moderating effect of expertise (number of reviews written), I fail to obtain the empirical evidence to support my hypothesis. One of the reason could be that "number of reviews written" is more a measure of 'experiences' rather than of 'expertise'. For example, a reviewer may have visits to many restaurants (i.e., an experienced reviewer) but still does not possess the superior knowledge repository and knowledge structure to evaluate restaurants (i.e., an expertise reviewer). In the future, I will modify the measurement of reviewer expertise to further validate my model.

In addition, I find that restaurant niche width magnifies the usefulness of the certainty-embedded review by a popular reviewer while mitigating the usefulness of the certainty-embedded review by an expert reviewer, thereby supporting my hypotheses. I conjecture that this circumstance can be attributed to the fact that the signal of large niche width (e.g., a restaurant that occupies multiple cuisines) increases the difficulty in evaluation; such an increase readily activates customers' systematic processing of the certainty-embedded review when they identify with a popular reviewer. By contrast, the small niche width (e.g., a restaurant that occupies only one single cuisine) symbolizes ease of understanding and tends to strengthen systematic processing when customers internalize with expert reviewers to absorb knowledge.

1.7.2 Theoretical Implication

Drawing on dual-process theory and social influence theory, I provide a theoretical model to understand the interactions among review certainty, reviewer expertise, reviewer popularity, and organization niche width on review usefulness. I ground the real-world count of usefulness votes in theory by linking it to the elaboration concept. My findings help extend the literature on online customer reviews in multiple aspects.

First, the majority of the prior research on review usefulness/helpfulness has focused on the easily observable determinants such as numeric ratings, review length, and reviewer reputation. Thus, only a few studies have investigated the textual content of reviews with the exception of Yin et al (2014). Addressing this research gap, my research adds to the merging body of the text mining literature (see, e.g., Kovács et al 2013; Humphreys 2010) by emphasizing the effects of the certainty embedded in review content on review usefulness. I document evidence of the adaptive nature of review certainty in review usefulness and report robust evidence on the important role of reviewer characteristics-e.g., popularity on the usefulness of the certainty-embedded review. The results clarify the mixed evidence in the literature on certainty. As I mentioned in the beginning of my study, Sniezek & Van Swol (2001) empirically demonstrated the positive effect of certainty on the information trustworthiness and acceptance while Yin et al (2014) indirectly presented the negative impact of the certainty embedded in emotions as one dimension of cognitive appraisal on the elaboration likelihood. These findings provide empirical evidence for a contingent model showing that, depending on reviewer popularity, the certainty embedded in reviews either strengthen (e.g., reviewer with a few fans) or weaken (e.g., reviewer with many fans) the usefulness of reviews. By doing so, I generate insights into the understanding of certainty impact, varying by reviewer popularity.

Second, my findings supplement the review usefulness/helpfulness literature (see, e.g., Yin et al 2014; Mudambi & Schuff 2010; Forman et al 2008) by identifying two reviewer characteristics—expertise and popularity and their interactions with review certainty. I find that a reviewer followed by many fans signals low usefulness of the certainty-embedded review. These findings imply the importance of match between review and reviewer in producing a useful review. Besides, my findings also extend dual-process theory by examining reviewer popularity, an additional peripheral cue that has received less attention in dual-process research but is commonly found in the context of social media.

Third, this study provides new insights into understanding the value of context by identifying a context-specific factor—i.e., organization niche width to formulate context-sensitive predictors of the persuasion of online customer reviews. Considering that social communications occur in different contexts, such as hotel, restaurant, groceries, etc., e-WOM has different implications for processing messages in different contexts. In my study, I focus on the restaurant context. In light of the specific restaurant contextual feature—i.e., cuisine niche width, I find that that restaurant niche width magnifies the usefulness of the certainty-embedded review by a popular reviewer while mitigating the usefulness of the certainty-embedded review by an expert reviewer. My findings address a call issued by (Hong et al 2013) to incorporate context into theory development and fit in their framework by considering contextual factors as moderators of proposed relationships.

1.7.3 Practical Implication

In the presence of uncertainty of product quality, consumers have to balance their private knowledge with the inferences drawn from opinions of predecessors who have consumed the products. Then online customer reviews have been treated as important references in decision making (Martin & Lueg 2013; Senecal & Nantel 2004), especially for useful reviews (Chen et al 2008). My study can offer some actionable implications for managers as well as retailers in the prediction and utilization of useful reviews.

First, my findings regarding the determinants of review usefulness can provide prominent benefits for online third-party review websites in the screening and selection of useful information. At present, the review voting system is post-hoc in which the most useful review can only be known after peer participation in voting action. However, my findings can enhance the ante-hoc mechanism in review recommendation system. For example, Yelp.com can consider the certainty factor in updating its default sorting mechanism—i.e., “Yelp sort,” by putting certainty-embedded reviews on the top of a business’s information page except for the fixed features (e.g., date, rating) to simplify user’s access to valuable reviews.

Second, my findings with respect to the moderating effects of organization niche width imply that restaurant managers should view online customer reviews as a double-edged sword and better provide different guidelines for customer review-writing depending on the expertise, popularity of the reviewer, and niche width of the target organization. For a restaurant occupying multiple cuisines—e.g., tagged with “Mediterranean”, “Greek” and “Middle Eastern”, the possible large variance of opinions embedded in reviews from different reviewers with various preferences can amplify the difficulty for potential customers to take a complete understanding of the restaurant. Taking the circumstance into account, managers can encourage more popular customers with many fans to write reviews for this restaurant, engendering the identification impacts and downstream purchase decisions. In contrast, for a restaurant solely serving a specific cuisine, managers better strengthen customers’ understanding of the restaurant deeply and completely by attracting more experts with professional knowledge and rich experience on the specific cuisine to write reviews for the restaurant.

Finally, at a broad level, my findings also give some managerial implications for social media platforms (e.g., Facebook, Twitter). With the emergence of social buttons in social media—e.g., ‘like’, ‘share’, ‘comment’ icons below a post on Facebook, users can interact with other people by clicking corresponding social buttons, as result, this activity can then turn into numbers on the associated button counter to signal the attractiveness of a post (Iturrioz et al 2014). Given the critical importance of customer engagement or interactivity in e-commerce, a frequently discussed question for managers when implementing social media marketing is how to reinforce social interactivity. My findings predicting what makes a useful review can also shed light on what generates a heated topic reflected by the total number of social buttons.

1.7.4 Limitations and Future Research

My study has a few limitations. First, the use of data from Yelp.com has the advantage of being an objective approach compared with other methods that capture subjective perceptions. Given that the controlled experimental method has the advantage of increasing internal validity, future experimental studies can provide converging evidence for my model. Second, text-mining techniques enable interesting areas for future research. Content analysis can be used to obtain additional information from review content to further explain what constitutes a useful review. Future research can examine other textual features of reviews, such as information richness and the possible interaction effects. Finally, my predictive model can also be extended to include the visual determinants of review usefulness, including the visual quality of pictures embedded in a review. At present, online customer reviews consist of not only textual content and numeric ratings but also visual pictures; such images are ubiquitous on online review platforms (e.g., Yelp.com). Future studies can extend the current predictive model to examine how the visual quality of attached pictures with reviews influences review usefulness.

For empirical testing, given the limited validation of the impact of reviewer expertise, in the future the measure of expertise should be refined—e.g., the number of useful votes received before he/she wrote the focal review, his/her past visits to and familiarity with similar restaurants, and his/her reputation, a reviewer's historical rating of reviews as useful. In this study, expertise here highlights reviewer's rich experience in writing reviews and professional knowledge in specific areas rather than his/her experience in rating others' reviews. In my opinion, the percentage of reviews rated as useful can suggest the extent of strictness of a reviewer evaluating the usefulness of review. That is, the more reviews as useful in a reviewer's historical rating, the less stricter the reviewer could be in rating reviews. However, the rating strictness could not fully depict a reviewer's expertise. In the future, I would like to try other measures of reviewer expertise to see other possibilities. Regarding text analysis, I would like to take into account other textual features such as double negatives and sarcasm in the future. In addition, for robustness checks using random sample split, I will conduct additional analysis of random subsample dataset in the future if possible. In addition, I only examined restaurant reviews; hence, the generalizability of my findings to other contexts demands further empirical studies. Future research must consider other business categories (e.g., hotel and beauty). Also, I only focused on restaurants in the United States. Considering the rapid growth of online customer reviews throughout the whole world, further validation of my hypotheses in other countries (e.g., China) would provide additional insights into cross-culture research.

Chapter 2

How Prior Users' Helpfulness Votes on a Review Influence Subsequent Users' Trust of the Review and Corresponding Product Evaluations in E-commerce Context

2.1 Introduction

People use voting as a means of making proximal decisions. For example, voting is practiced in political election to select the most promising candidate based on the voting inclination of the majority. The recent concept of voting indicates the limited access of participants to voting dynamics without knowledge of prior votes before publicity of final voting results. By contrast, modern online voting systems highlight the transparency and real-time dissemination of the entire voting process. This process suggests that users may use the votes of prior users as reference for their own judgment. With the enforcement of user-to-user online interactivity in modern online settings, users can publicly express their opinions as well as voting others' opinions. Opinion voting implemented by webmasters takes place in widely used participatory platforms, ranging from social media sites to third-party review websites. Helpfulness voting system is a typical form of online voting system embedded in online review websites (e.g., Amazon.com, Dine.com). This system can be used by users to publicly evaluate the helpfulness of a review, which not only presents product reviews contributed by users, but also displays votes by other users on the review. Given that voting systems play an important role in detecting helpful reviews, helpfulness voting system has gained intensive attention from academics, particularly on the determinants of helpfulness votes on a review (Mudambi & Schuff 2010; Yin et al. 2014; Cao et al. 2011). These studies aim to answer the following question: What makes a helpful review?

However, a fundamental question should be clarified before exploiting the predictability of the helpfulness voting system. What is the outcome of this voting system for online sites that enable the publicity and dissemination of user opinions on a product or service? Given the exposure of previous opinions to potential users, I go further on the voting system by investigating the impact of votes of prior users on the decision-making of other users. Instead of asking "What did you think of the review?" I ask "What would subsequent users think of the votes by prior users on the review?" I explore whether the helpfulness voting system also works for online shopping behavior in addition to its original function of detecting helpful reviews. I propose that the votes contributed by prior users exert social influence on subsequent users. They tend to rely on prior users' votes as a rational act of trust building on the review of others user, and the corresponding attitude towards the product/service discussed in the review.

Systematic rules are not observed on the design of the voting system because of the availability of the helpfulness voting system across sites. For example, Amazon.com demonstrates the helpfulness of reviews as "10 people found the review helpful," which is a response to the question: "Was this review helpful to you?" Dine.com presents both the number of helpful and unhelpful votes beside the "YES" and "NO" counters. Two key metrics of prior votes are consistently demonstrated, namely, ratio and magnitude, regardless of the distinct displays of user-generated votes. Ratio indicates the proportion between helpful votes and unhelpful votes, and magnitude depicts the total number of votes, including both helpful and unhelpful votes. Thus, helpfulness of a review can be interpreted by users referring to either the ratio or magnitude. This approach raises a question of whether subsequent users may react differently or selectively to the ratio or magnitude of prior users' votes. The response depends on how the metrics are used in the review's elaboration and internalization. Therefore, I propose a two-stage model that illustrates the sequence of processing helpfulness votes. In this sequence, ratio, as a primitive and easy-to-interpret metric, is utilized first. Magnitude, as a less diagnostic and calculation-demanding metric, is used subsequently on the conditions of relative motivation and ability.

I provide empirical evidence to support my model. Three experiments are conducted in this study to test the link between helpfulness votes and user attitude towards the product/service discussed in the review. The results showed that the helpfulness voting system plays a role in the decision-making process of users. Regardless of the valence and type of reviews, high helpfulness ratio enhances their trustworthiness and guide corresponding behavior. Next, attention turns to the effect of vote magnitude. Unlike vote ratio, which is applicable to both positive and negative reviews, helpfulness magnitude is only significantly influential for negative reviews, which motivate users to take additional cues for decision making. Despite the degree of helpfulness ratio, users will avail helpfulness magnitude to evaluate the review and make a corresponding decision. Furthermore, unlike the helpfulness ratio, which is applicable to both attribute- review and emotion-based reviews, helpfulness magnitude is only significantly influential for attribute-based review, which offers users the ability to take a rational reference of the prior votes for their product evaluations and purchase intentions. Hence, users will take advantage of additional helpfulness votes as a ratio constant when the review is untrustworthy. In a word, my results suggest that the processing of helpfulness votes follows a sequence, i.e., ratio will be utilized first and magnitude will be used later when users have the motivation and ability to internalize it as reference in decision making.

My findings offer several contributions. First, I extend online product review literature by presenting a comprehensive way of conceptualizing helpfulness votes and review effectiveness. Most existing studies on review helpfulness aim to identify the determinants of review helpfulness votes, including the content characteristics, which include review length (Mudambi & Schuff 2010) and emotions (Yin et al. 2014), and source characteristics, which include self-identity disclosure (Forman et al. 2008); however, the consequences of helpfulness votes, such as its effect on user's attitude, are not explored sufficiently. The present study is the first to investigate the effectiveness of helpfulness voting system on the online behavior of users. Unlike existing studies that examined the link between Facebook likes and product sales (Kuan et al. 2014; Lee et al. 2015), which suggests how product endorsements temper a user's purchase intention, the present study exploits how prior users' votes on the review tease the purchase of another user on a recommended product in the review. Instead of simply focusing on the number of Facebook Likes (FBLs) (Kuan et al. 2014), I consider the numerical ratio of helpfulness votes as a means of examining the valence of social influence, thereby extending previous studies that manipulated the valence of social influence by self-reports (Graziano et al. 1993) or separately regarded up-votes as positive influence and down-votes as negative influence (Muchnik et al. 2013). Unlike existing studies that focused on the voting behavior induced by prior votes (e.g., Muchnik et al. 2013), which reveals herding effect in voting behavior, the present study extends the social influence exerted by prior votes from its effect on users' voting behavior to how subsequent users utilize votes for assessing review and the corresponding product discussed in the review.

Also, this research has important managerial implications. My findings suggest that beyond the detection of helpful reviews, the helpfulness voting system also plays a role in the online users' decision-making process by enabling them to obtain additional information and make a rational evaluation of reviewed product or service. Online product reviews amplify and accelerate the reach of marketers to the point that nearly any user feedback on products or services can function as an influential information source. In addition to the content of reviews, my findings identify another means, namely, helpfulness votes, through which online retailers can increase the influence of their retail sites.

The rest of this study is organized as follows. First, I illustrate the theoretical background of the study by conceptualizing prior votes as normative influence and the review content as informational influence. I propose a set of hypotheses on the effects of votes and their interactions with review characteristics based on the framework. I then offer an overview of all studies and present my empirical results. Finally, I conclude the research and provide future research directions.

2.2 Literature Review

2.2.1 Online Product Reviews

As one form of electronic word-of-mouth, online product reviews, which refer to informational communications among customers concerning evaluations of goods and services, have played an increasingly important role in electronic commerce in various aspects, such as informing potential customers of product knowledge (Martin & Lueg 2013), reducing uncertainty in product quality (Senecal & Nantel 2004), and increasing product sales (Chevalier & Mayzlin 2006). To the best of my knowledge, two typical dimensions of online customer reviews have been frequently investigated, namely, valence and type.

Concerning review valence, beyond measuring it using five-star ratings, an emerging stream of studies have deployed text-mining approach to measure textual review valence in continuum (e.g., Ludwig et al. 2013; Goh et al. 2013; Tirunillai & Tellis 2012). For example, Ludwig et al. (2013) revealed a quadratic relationship between changes in affective content and changes in buying. Tirunillai & Tellis (2012) used sentiment analysis and found that negative content exerts stronger influence on returns than positive content does. Despite a large number of studies demonstrating the significant effect of review valence on economic outcomes, such as purchase decisions and product sales (e.g., Chevalier & Mayzlin 2006), several studies found insignificant empirical effect of valence (Duan et al. 2009; Liu 2006), and a few studies reported mixed results for valence (e.g., Berger et al. 2010; Ludwig et al. 2013).

In terms of review type, Holbrook (1978) suggested two kinds of persuasion evidence, namely, factual evidence defined as "logical, objectively verifiable descriptions of tangible product features" and evaluative evidence defined as "emotional, subjective impressions of intangible aspects of the product." A few studies examined the impact of factual versus evaluative evidence in a review on consumer behavior (Hong et al. 2012; Lee et al. 2008; Park et al. 2009); however, these studies did not give a thorough definition of such evidence. For example, Lee et al. (2008) presumed that evaluative review without providing product-related information was of low quality, whereas factual review listing reasons supporting argument was of high quality. Park et al. (2009) further defined a review mostly consisting of factual evidence as attribute value review and a review mostly made up of only evaluative evidence as simple recommendation review. Thus, the impacts of these two types of review content should be differentiated in a more cautious way.

2.2.2 Review Helpfulness

Concerning research on review helpfulness, most of current studies aim to identify the determinants of review helpfulness (e.g., Mudambi & Schuff 2010; Yin et al. 2014; Cao et al. 2011), and inconsiderable research directly studied the numeric characteristics of helpfulness information and their consequences, except for some studies that investigated the mediation role of perceived review helpfulness (e.g., Purnawirawan et al. 2012).

Existing studies on the determinants of review helpfulness (e.g., Yin et al. 2014; Mudambi & Schuff 2010; Forman et al. 2008) mainly examined two key components of an online customer review, namely, review and reviewer. One stream of studies examined numeric ratings and found that negative reviews are more useful in customer decision making than positive reviews are. For instance, Mudambi & Schuff (2010) examined the factors influencing the perceived review helpfulness from the information diagnosticity perspective, asserting that rating extremity influenced review helpfulness. In recent years, scholars directly investigated review text using text-mining techniques. For instance, Cao et al. (2011) applied a text-mining technique to systematically examine the basic, stylistic, and semantic characteristics of reviews using a unique dataset of CNET.com software reviews. Beyond linguistic characteristics, Yin et al. (2014) showed that the emotions embedded in a review influenced the perception of customers about review helpfulness and therefore proposed that anxiety-embedded reviews were more helpful than anger-embedded reviews were via the mechanism of perceived cognitive effort. Aside from review content factors, the characteristics of reviewers, such as reviewer authorship (e.g., Forman et al. 2008) and reviewer reputation (e.g., Otterbacher 2009) also influence review helpfulness.

With respect to the impacts of review helpfulness, a few relative studies exist. For instance, Purnawirawan et al. (2012) demonstrated the association between perceived review helpfulness and customer attitude, indirectly implying the importance of helpfulness of reviews. However, perceived review helpfulness is not a good way to illustrate the numeric characteristics of review helpfulness in online review platforms (e.g., Amazon.com). Chen et al. (2008) directly identified a larger impact of reviews with a high proportion of helpful votes for less popular books. However, both studies do not give a complete and direct examination of helpfulness information itself. Thus, more direct and comprehensive depiction of the helpfulness voting system is needed, as well as more empirical evidence for the effectiveness of review helpfulness votes.

2.3 Theoretical Background and Conceptual Model

2.3.1 User Trust and Social Influence

Given the information asymmetry between buyer and seller (Ackerloff 1970), a buyer has to search for trustworthy information as reference for his or her product evaluations and final purchase decisions. Customers who have uncertain purchase decisions can rely on the observed collective opinion or action polarity to reduce uncertainty and enhance trust in their decisions. Trust is demonstrated as important antecedents of behaviors that rely on the advice or actions of others (Mayer et al. 1995). Research about information search suggests that following others can reduce the time and energy associated with searching information and experimentation (Muchnik et al. 2013). Therefore, I expect that social influence in the form of online product reviews will function as trust cues to assist user evaluations.

Social influence plays an important role in the consumption process (Burnkrant & Cousineau 1975). Social influence reflects the perception of an individual toward the behavior influence of others. Social influence reflects individuals' perception of others' behavior influence. That is, the observed collective behaviors or opinion polarity can be regarded as information such as high-quality signal, to help make a decision. On the other hand, the observed behaviors can be perceived as normative pressure, that is, the decision-making is not owing to the obtainment of information by observation but because the group pressure. In a word, others' behavior can be perceived as either a kind of information for usage called informational influence or the group pressure that have to obey called normative influence (Deutsch & Gerard 1955). According to Kelman (1961), social influence operates through three distinct processes termed compliance, identification, and internalization. Internalization occurs when people accept the majority opinion and integrate it into their belief systems. Compliance occurs when people conform to the majority opinion but obtain their original opinions outside the group influence situation. Identification arises when people adopt group attitudes or behaviors to establish a relationship with the group. These three processes can relate to Deutsch and Gerard (1955)'s two forms of social influence, informational influence and normative influence. Informational influence leads to acceptance of information, which corresponds to internalization while normative influence is the tendency to conform to the expectations of others, which could be attributed to compliance or identification.

2.3.2 Two-stage Model

Given that review votes are accumulated over time and may undergo systematic changes through time, votes on a review may fluctuate before reaching final consensus. Thus, users may come across helpfulness votes in different ratios or magnitudes. For example, user A at time point A may read the review with 30 helpful votes and 4 unhelpful votes, which implies that most predecessors perceive the review as helpful. However, user B at time point B may see the review with 30 helpful votes and 30 unhelpful votes, which implies that the evaluation of helpfulness of the review is in conflict without coming to consensus. Moreover, user C at time point C may find the review with 300 helpful votes and 40 unhelpful votes, which implies an increasing number of users voting the review.

I propose a two-stage model to understand how the review helpfulness numbers generated by prior users influence the decision of subsequent users. In this study, I propose that the selection of cues also occurs between heuristic cues. I assert that the processing of heuristic cues occurs in sequence, which means that easy-to-understand cues will be utilized first. Less diagnostic cues will be used later, when users have motivation and ability to internalize the cues.

Users self-select heuristic cues for information internationalization and downstream evaluations. Thus, I assert that customers will first utilize the fundamental cue, which is diagnostic. Then, they need limited cognitive effort consumption or the helpfulness ratio to decide whether they will follow the review or not. This situation is called the first stage. After users decide on following the review, they will use the secondary cue or the helpfulness magnitude for further elaboration, especially when they think additional efforts are necessary. From the perspective of conformity effect induced by review helpfulness votes, my two-stage model shows that vote ratio (high helpfulness ratio vs. low helpfulness ratio) can explain the direction of conformity, namely, follow or deny. By contrast, vote magnitude (small vs. large) can explain the strength of conformity, namely, strong or weak following.

Because ratio and magnitude appear simultaneously, in reality users' decision on the trust of a review is made based on the overall consideration of both ratio and magnitude. But I propose that the processing of the two metrics follow a sequence in users' mind. In order to uncover the underlying decision-making process induced by votes, I separately study the impacts of ratio and magnitude and tell their differences. To clarify the sequence of self-selection in the ratio and magnitude cues, I argue the unconditional use of ratio cue. The implementation of magnitude cue requires elaboration motivation (e.g., negative review) and elaboration ability (e.g., attribute-based review). When customers have the motivation or ability to elaborate further the review, they will use magnitude as reference. Moreover, small magnitude suggests low user engagement. It is possible that users are likely to pay little attention to the review with a small number of votes. Following this logic, magnitude works more like the attention-capturer rather than trust-builder, because it rarely signals the trustworthiness without the hints of vote ratio. In other words, if magnitude goes first in decision tree, users are still uncertain of following the review or not if they don't know the percentage of up-voting to the total number of votes. I hypothesize that at first ratio can inform users of following or not following and then they refer to magnitude and review content in determining to what extent they will follow the review. Empirical studies suggest that elaboration ability and motivation are two key factors that influence the likelihood of elaboration (Angst & Agarwal 2009).

As shown in Figure2, the logic depicting the ratio impact is that individuals in nature follow/trust the majority. Regarding vote ratio, when the majority of individuals think of the review as helpful, i.e., high helpfulness ratio, subsequent customers are likely to follow the majority to adopt the review, and then trust the review more. By contrast, for the low helpfulness ratio with the majority people viewing the review as unhelpful, subsequent customers are supposed to follow the majority opinion and not to trust the review.

With respect to the magnitude, as the ratio is constant, the increase of total votes indicates the increase of the difference between the helpfulness and unhelpfulness votes, suggesting the salience of the majority relative to the minority. Concretely speaking, for the large magnitude, viewers' inclination of following the majority could be enhanced, thus magnifying the trustworthiness of the review. In contrast, when the magnitude is small with a small win between up-voting and down-voting, viewers are likely to hesitate and swing in trusting the review.

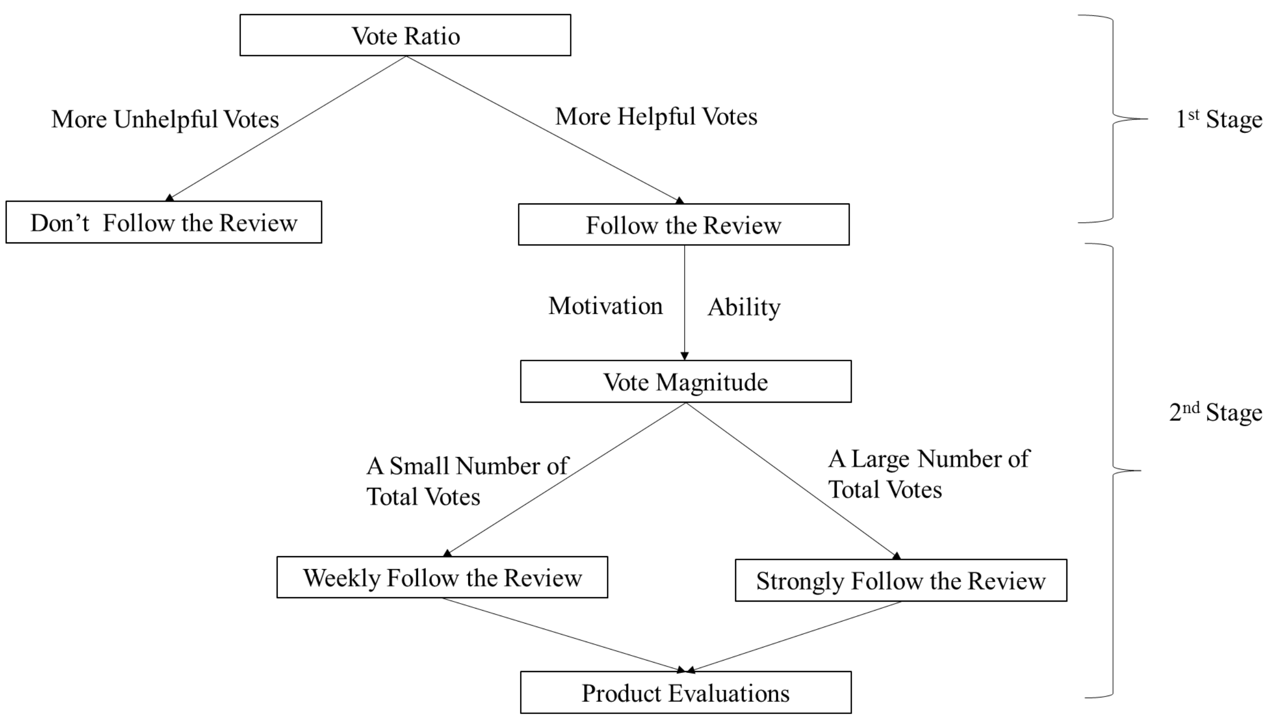


Figure 2. Decision Tree: The Influence of Review Helpfulness Votes on the Utilization of Online Product Review

2.4 Research Model and Hypotheses

2.4.1 Research Model

I draw on social influence theory to explain the underlying rationales behind the research model (see Figure 3). Deutsch and Gerard (1955) proposed two forms of social influence, namely, informational and normative influence. Informational influence is based on the acceptance of information from others as evidence of reality. This study viewed review content as informational influence. This study focused on the two aspects of review content, namely, valence and type. Concerning valence, I study the two polarities, positive and negative affect. Examples of positive reviews, broadly defined, relate to pleasant, useful, or interesting experiences, while negative review includes unpleasant experiences or complaints on the product. In terms of type, I propose differential impact of two types of review, which are attribute- and emotion-based. Attribute-based review indicates reasonable evaluation based on tangible features of a product, while an emotion-based review is the emotional evaluation without referring to tangible aspects of the product.

Normative influence is the tendency to conform to expectations of others. Ratio indicates the direction of normative influence, that is, the positive influence delivering agreement, whereas negative influence communicates disagreement (Muchnik et al. 2013). Muchnik et al. (2013) claimed that the conformity of positive social influence with up-vote counts more than down-vote counts, whereas the correction effect for negative social influence with up-vote counts less than down-vote counts. By applying the valence of social influence in review helpfulness, Vote ratio implies the percentage of peers viewing the review as helpful, high helpfulness ratio, i.e., dominating more helpful votes than unhelpful votes, low helpfulness ratio, i.e., more unhelpful votes than helpful votes. Magnitude indicates the strength of normative influence, which is the strong conformity indicating the large size of the majority, whereas weak social influence is related to a small size of the majority. Vote magnitude indicates the total number of votes, including both helpful and unhelpful votes.

As shown in Figure 3, I illustrate four factors to explain user trust and attitude, namely, vote ratio and magnitude, as well as review type and valence. As discussed above, Informational influence is to depict review content while normative influence is to reflect the impact of vote ratio and magnitude. The main objective of this study is to validate the decision tree, that is, vote ratio acts as an unconditional trust-cue while vote magnitude functions as a conditional trust-cue, dependent on users' motivation and ability manifested in review content. Thus, vote ratio and magnitude are my focus to be theorized to directly influence users' trust of review and attitude. Review content is used to demonstrate the conditions on which vote ratio and magnitude are effective. Moreover, the underlying logic of the research model is shaped in such a way that vote ratio and magnitude are expected to influence users' trust of the review, and the more trust endowed on the review, the greater internalization of the opinions in the review will be, and finally the more corresponding attitude will be developed. The preceding argumentation goes in such a way that people follow the majority and accordingly build up trust, and finally refer to opinions in the review for decision-making. In other words, I propose the existence of herd behavior when encountering predecessors' votes. In terms of the moderating role of review type, I further investigate the herd behavior induced by previous votes and emphasize that what I focus on is rational herding rather than irrational/mindless herding. That's why I examine the impacts of vote ratio varying by review type, informative attribute-based review vs. less informative emotion-based review. Thus, I argue that vote ratio and magnitude lead to different degrees of trust and attitude, depending on review type and valence.

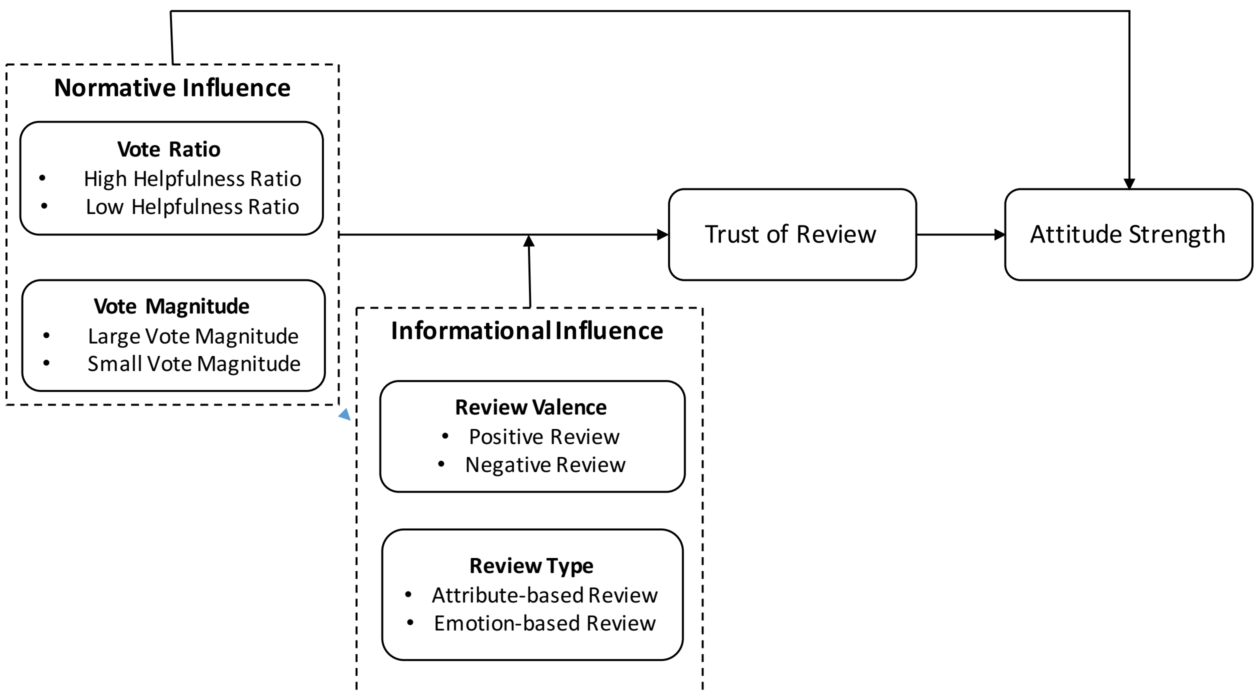


Figure 3. Model of Vote Ratio and Magnitude

2.4.2 Impact of Vote Ratio

In terms of high helpfulness ratio, a substantially high number of peers who view the review as helpful indicate that the review content can be understood, and most of the peers agreed with the reviewer's comments on product or service. The majority of previous viewers of the review who reached an agreement on the helpfulness of the review indicate that the reviewer tells the truth that the recommended product or service is as good (bad) as described in the positive (negative) review, which strengthens the credibility of the review. Therefore, high helpfulness ratio enhances users' trust of the review. By contrast, low helpfulness ratio indicates that majority of previous viewers of the review think that the review is not helpful. This review may be composed of complicated or irrelevant words, which are difficult to comprehend and understand, and high complexity or readability feature (Cao et al. 2011). The high number of unhelpful votes is attributed to low quality of the review itself. Thus, customers experience difficulty in inferring the quality of the review, thereby decreasing their trust on the review and conspicuousness about product quality discussed in the review.

According to prospect theory, positive reviews induce an adoption of the reference point. The corresponding decision is perceived as possible gains, whereas in negative reviews, the purchase decision is perceived as possible losses and may elicit avoidance of such behavior (Kahneman & Tversky 1979). Therefore, when the majority of individuals think of the review as helpful (unhelpful), i.e., high (low) helpfulness ratio, subsequent customers are likely to follow the majority to adopt (ignore) the review, and are likely (unlikely) to favor (disfavor) the product in the positive (negative) review to others. A high review helpfulness ratio with the majority perceiving the review as helpful is more likely to strengthen the attitude toward the product mentioned in the review to others; conversely, a low review helpfulness ratio has the majority perceiving the review as unhelpful.

H1a. Users exposed to a positive (negative) review with a high helpfulness ratio will have more favorable (unfavorable) attitude towards the evaluated product or service than those exposed to the review with a low ratio.

H1b. The impact of helpfulness ratio on users' attitude is mediated by their trust of the review.

2.4.3 Moderating Role of Review Type

I consider review type when examining the helpfulness ratio effect. Here, I propose two types of reviews, namely, attribute- and emotion-based. Attribute-based review indicates reasonable evaluations based on tangible features of a product, whereas emotion-based review is the emotional evaluation without referring to the tangible aspects of the product. For example, one review, i.e., "I can't believe I got it. I'm proud of it," is subjective, emotional, and has no support for arguments. This type of review is emotion-based. By contrast, an attribute-based review provides comments on product-related benefits, such as, "This restaurant provides good service and delicious food," which is specific, clear, and offers reasons for favorable recommendations.

Online users are more rational in decision making or conduct rational herding (Zhang & Liu 2012). That is to say, online users will not completely follow others without any basic information. Zhang and Liu (2012) suggest that rational herding, which is active observational learning, beats down irrational herding, which is passively mimicking predecessors. Considering that one following the majority is based on the assumption that the person takes the fact for granted, and that the majority group makes the decision based on a complete understanding or possession of relative information (Anderson & Holt 1997). Thus, I propose the next hypothesis:

H2. Review type moderates the vote ratio effect such that the vote ratio effect is stronger when the review is attribute-based review than when it is the emotion-based review.

2.4.4 Impact of Vote Magnitude

Given the vote ratio constant, the large magnitude indicates the large size of the majority, whereas the small magnitude suggests the small size of the majority. Increasing the size of the majority implies the addition of users with the same view, which will further confirm the majority as a representative sample of the whole population. In terms of large magnitude, a large number of users reach consensus on the helpfulness of the review. The increasing size of the majority implicates underlying information about reality and strengthens the power of the majority to reward and punish. Moreover, large helpfulness magnitude implies a large difference between the number of helpful and unhelpful votes, which is dominating consensus. A review with a large vote magnitude is less uncertain and more trustworthy. By contrast, a low number of viewers for a small magnitude reach consensus on the helpfulness of the review. The small size of the majority indicates that the review may not be attractive and convincing to be adopted by many individuals because of the minor difference between the number of helpful and unhelpful votes, which is weak consensus.

2.4.5 Moderating Role of Review Valence and Type

According to prospect theory, individuals under risk interpret outcomes as gains and losses, and are more sensitive to losses than to commensurate gains or loss aversion (Kahneman & Tversky 1979). This behavior is attributed to individual experience of loss, which appears greater than the gains associated with obtaining an amount equivalent to that which was lost because the value function is steeper for losses than for gains (Kahneman & Tversky 1979). In this sense, users are supposed to be more sensitive to negative reviews, and thus need more cues to support their corresponding unfavorable evaluations than positive reviews. Thus, negative reviews are more likely to motivate users to take advantage of vote magnitude as reference for internalization of online negative information.

In addition, visible and clear signals help customers reduce their information search and processing costs. Therefore, effective information signals must be visible and clear (Rao & Monroe 1989). An emotion-based review that provides entire affective recommendation of a product is ambiguous, which may increase the information processing difficulty. By contrast, attribute-based reviews that consist of explicit product-related evaluations easily facilitate understanding. With regard to internalization process, the attribute-based review embedded with concrete information is clearly more beneficial or incremental to complete a task or resolve a problem. That is, the attribute-based review whose evaluations on various aspects of a product/service can be used as knowledge to be internalized, as the personal belief of an individual. By contrast, the emotion-based review that merely transfers abstract emotions can hardly increase the knowledge or upgrade the beliefs of the customers, let alone internalization. Thus, compared to emotion-based review, the attribute-based review enables users to make sense of the large helpfulness magnitude by stating concrete signals of product quality. I propose the next hypotheses.

H3a. As the vote ratio constant, review type, and review valence simultaneously moderate the magnitude effect such that only for the negative and attribute-based review, users exposed to a review with large vote magnitude will have lower attitude toward the evaluated product or service than those exposed to the review with a small magnitude.

H3b. The impact of vote magnitude on users' attitude is mediated by their trust of review.

2.5 Empirical Analysis

I presented across three studies participants with product decision scenarios that included information on evaluation on the product and the numerical judgment of prior customers on the evaluation in the form of textual reviews and prior users' helpful and unhelpful votes on a review. To provide a thorough illustration of the social influence induced by prior customers, Experiment

1 first demonstrated the generality of vote ratio effect by focusing on positive and negative reviews. I then examined the underlying process by measuring customer trust of review. If users utilize the high helpfulness ratio, they will trust the review and internalize the information embedded in the positive (negative) review as reference for product evaluation and downstream purchase decisions. By contrast, if customers internalize the low ratio as reference, they will belittle the review and lower the trust thereby inhibiting the subsequent outcomes as a result. In Experiment 2, I used positive review as the product evaluation. I then examined the emotion-based review to present further the generality of the ratio effect across review types.

Given the validation of helpfulness ratio effect based on Experiments 1 and 2, I addressed another metric, i.e., vote magnitude in Experiment 3. To tease out the magnitude effect, I focused on the helpful review varying with different degrees of vote magnitude. I focused on how the increase of helpfulness votes changes users' perception of a review and their corresponding behavior when the helpfulness ratio is constant. I found that magnitude is a conditional cue, which suggests that users need motivation and ability for taking advantage of vote magnitude for product evaluations.

2.5.1 Experiment 1

Experiment 1 had three main objectives. First, the setup tested the helpfulness ratio effect on attitude and trust. Two levels of ratio were examined, i.e., low ratio vs. high ratio. Second, I tested whether the ratio effect can be generalized to positive and negative reviews. Third, I tested the proposed underlying mechanism, that is, the helpfulness ratio affecting user attitude by strengthening or weakening their trust of the review.

2.5.1.1 Method

I recruited 203 participants (116 males; Mean age=35) from Amazon Mechanical Turk. Each participant was compensated for her or his time with 1US \$. The participants were randomly assigned to one of four conditions in a 2 (vote ratio: high helpfulness ratio vs. low helpfulness ratio) × 2 (review valence: positive vs. negative) between-subjects factorial design.

Initially, participants were asked to imagine the following scenario, "You find a restaurant. Before going to the restaurant, you look it up on a review website." They were subsequently exposed to a restaurant review. In the positive review condition, the participants were asked to read a review that recommended the restaurant by describing the positive features of the restaurant, including food, service, an environment. In the negative review condition, the participants were asked to read a review that cited unsatisfactory aspects of the restaurant. To manipulate the helpfulness ratio of the review, the following question was asked at the end of review, “Was this review helpful to you?” Varying “Yes” and “No” responses were recorded. In the high ratio condition, the participants will encounter 50 “Yes” and 6 “No,” whereas in the small ratio condition, they will face 6 “Yes” and 50 “No.” Participants were asked to answer several questions.

When participants finish reading the review, they were asked to answer questions with the following instruction, “Please answer the following questions based on the review you read.” At first, participants completed a three-item measure of attitude toward the restaurant ($M=3.972$, $SD=2.251$). The measure was adapted from Rucker and Petty (2004) by asking a question, “What do you think of this restaurant?” Items were answered on a seven-point scale (1=Bad/Unfavorable/Dislike, 7=Good/Favorable/Like; $\alpha=0.991$). Next, the participants completed the two-item measure of trust of the review ($M=4.599$, $SD=1.591$), i.e., “I trust this review” and “I can rely on this review.” Items were answered on a seven-point scale (1=Strongly disagree, 7=Strongly agree; $\alpha=0.958$).

3.5.1.2 Results

Manipulation Check. For the manipulation check of review valence, participants were asked to indicate the extent to which they agree with the review ranging from very negative to very positive on a seven-point scale (1=very negative, 7=very positive). The one-way ANOVA result indicated that the manipulation of review valence was successful ($F(1,201)=1424.277$, $p<0.001$) in such a way that participants in the positive review condition considered the review as positive ($M=6.304$, $SD=1.097$) while participants in the negative review condition considered the review as negative ($M=1.287$, $SD=0.766$).

Attitude. I found significant interaction effects of helpfulness ratio and review valence on attitude ($F(1,199)=19.983$, $p<0.001$), supporting H1a. For positive review, participants had more positive attitude toward the restaurant when reading the review with high helpfulness ratio ($M=6.267$, $SD=0.797$) than the review with low helpfulness ratio ($M=5.481$, $SD=1.395$; $F(1,199)=11.947$, $p=0.001<0.01$). For negative review, participants had less favorable attitude when reading the review with high helpfulness ratio (M=1.715, SD=0.917) than the negative review with low helpfulness ratio (M=2.371, SD=1.333; F(1,199)=8.220, p=0.005<0.01). Figure 4 illustrates the plot of the interaction effect.

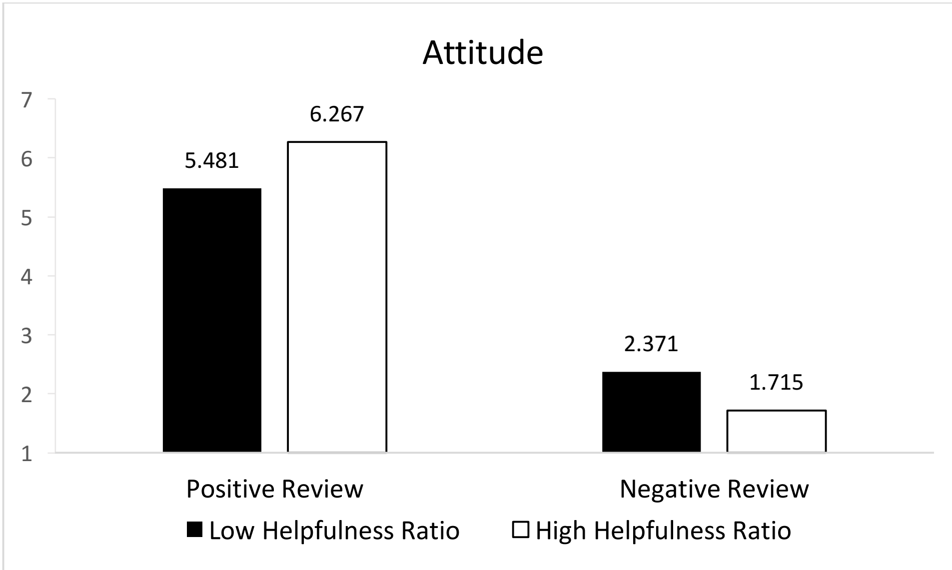


Figure 4. Impact of Helpfulness Ratio on Attitude across the Positive and Negative Review

Trust. I conducted a two-way ANOVA and obtained significant main effect of helpfulness ratio on trust (F (1,199)=16.358, p<0.001). The result did not indicate significant interaction effect of ratio and valence on trust (F(1,199)=0.499, p=0.481), which suggested that the helpfulness ratio effect was significantly influential for positive and negative reviews. For positive review, participants had higher trust of the review with high helpfulness ratio (M=5.290, SD=1.143) than the review with low helpfulness ratio (M=4.269, SD=1.682; F(1,199)=11.354, p=0.001<0.01). Likewise, for the negative review, participants had significantly higher trust with high helpfulness ratio (M=4.802, SD=1.529) than the negative review with low helpfulness ratio (M=4.085, SD=1.683; F(1,199)=5.538, p=0.020<0.05). The plot of the ratio effect across the positive and negative reviews on trust is indicated in Figure 5 below.

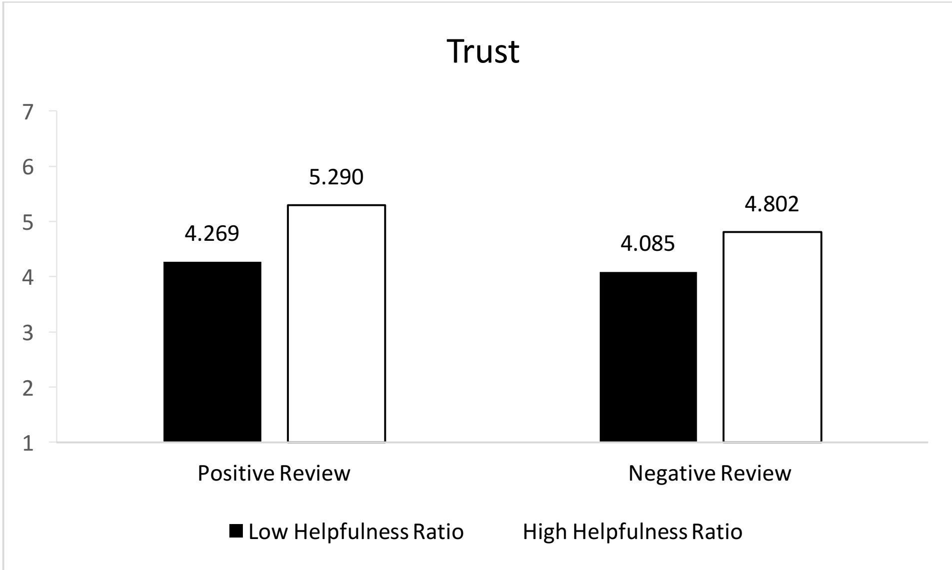


Figure 5. Impact of Helpfulness Ratio on Trust across Positive and Negative Reviews

Mediation Analysis. I applied Process Model 15 (Hayes 2013) and obtained the following results. The mediation effect of trust on the relationship between helpfulness ratio and attitude (95% confidence interval: 0.4747 to 1.4673) is salient across positive and negative reviews, supporting H1b. Furthermore, for the positive review, trust positively mediated the effect of helpfulness ratio on attitude (95% confidence interval: 0.2586 to 0.8397). For the negative review, trust negatively mediated the influence of helpfulness ratio on attitude (95% confidence interval: -0.6744 to -0.2093).

2.5.1.3 Conclusion

Experiment 1 demonstrated that helpfulness ratio can act as a trust cue that directly influences user attitude. Second, results suggest that compared with the review with low helpfulness ratio, a positive review with a high helpfulness ratio is more influential in fostering favorable attitude of users. By contrast, a negative review with high helpfulness ratio is more influential in driving unfavorable attitude.

2.5.2 Experiment 2

Experiment 2 had three main objectives. First, I further tested the boundary of the role of helpfulness ratio as trust cue by examining its application in different types of reviews. I examined whether the ratio effect was also salient for emotion-based review. Second, unlike

Experiment 1, which used 50 as the number of helpful votes with zero as digit, I set the high helpfulness ratio as 54:2 and low helpfulness ratio as 2:54. All numbers have non-zero digits. Third, unlike Experiment 1 that used textual "Yes" and "No" to represent user opinions on review helpfulness, this experiment manipulated "helpful" as the "thumb-up" sign and "unhelpful" as the "thumb-down" sign. Changing the number and the signal for manipulations can increase the generality of my vote ratio effect.

Experiment 1 demonstrated that ratio effect existed in both positive and negative reviews. Positive reviews are prevalent on online review sites. Thus, Experiment 2 focused on the positive review to further test the moderating role of review type on the vote ratio effect.

2.5.2.1 Method

A total of 86 undergraduate students from a medium-sized University in Hong Kong (22 males, mean age=22 years) were randomly assigned to one of four conditions in a 2 (vote ratio: high vs. low) × 2 (review type: attribute-based vs. emotion-based) between-subject factorial design. Each participant was compensated for her or his time with 1 Macdonald Coupon priced 10 HK \$. The reasons for choosing student sample are as follows: First, college students often browse the Internet and purchase products online. Second, this study requires a large sample size and students are the most accessible population. Third, given their limited knowledge on judging product quality and their willingness to listen to the opinions of others, college students depend on online customer reviews when making purchase decisions. Therefore, the perceptions of college students toward these reviews can provide valuable insights for my study. Finally, I have little reason to believe that the decisions of students influenced by online customer reviews are different from the decisions of other people because human decisions result from the collection and transmission of information into cognitive and behavioral systems (Panksepp 2005).

The procedures are the same as that in Experiment 1. In the attribute-based condition, the participants read a review that commented on various aspects of the restaurant, including food, service, and environment. In the emotion-based condition, participants read a review that merely expresses user's emotions about the restaurant without providing any reason to support the comments. The ratios of review helpfulness are high, and low. In the high helpfulness ratio condition, the participants are exposed to 54 "likes" and 2 "dislikes" of the review. In the low helpfulness ratio condition, the participants are exposed to 2 "likes" and 54 "dislikes" of the review. The participants completed a three-item measure of attitude toward the restaurant (M=4.163, SD=1.274), which was adapted from Rucker and Petty (2004) by asking the question, "What do you think of this restaurant?" The items were answered on a seven-point scale (1=Bad/Unfavorable/Dislike, 7=Good/Favorable/Like; **α**=0.954).

3.5.2.2 Results

Manipulation Check. To verify whether the two types of positive reviews were perceived as intended, participants provided answers related to the extent by which they agreed with the statement, "The review tells me the features of the restaurant (restaurant features include food, service, convenience of location, etc.)" on a seven-point Likert scale (from 1 = "strongly disagree" to 7 = "strongly agree"). One-way ANOVA yielded a significant main effect of review type (F (1, 84) = 34.314, p<0.001) in such a way that the participants in the attribute-based condition (M=4.023, SD=1.566) reported more than the participants in the emotion-based condition did (M=2.116, SD=1.451). This finding suggests that the manipulation of review type was successful.

Attitude. A two-way ANOVA on the attitude toward the product indicated significant main effect of helpfulness ratio (F(1,82)=36.151, p<0.001) and significant moderating effect of review type (F (1,82)=9.488, p=0.003< 0.01), thereby supporting H1b. For attribute-based review, participants had more favorable attitude toward the restaurant for the review with high helpfulness ratio (M=5.197, SD=0.710) than the review with low helpfulness ratio (M=3.159, SD=1.508; F(1,82)=41.341, p<0.001). For emotion-based review, participants also had more favorable attitude toward the restaurant when reading the review with high helpfulness ratio (M=4.460, SD=0.771) than the review with low helpfulness ratio (M=3.803, SD=0.990; F (1,82)=4.299, p=0.041<0.05). The plot of the interaction effect is indicated in Figure 6 below.

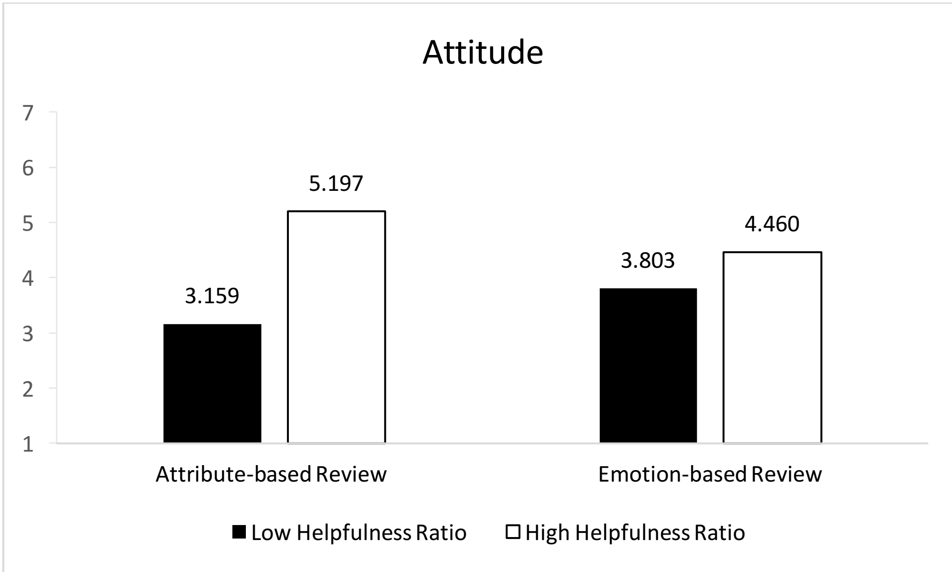


Figure 6. Interaction Impact of Helpfulness Ratio and Review Type on Attitude

2.5.2.3 Conclusion

These findings support my hypotheses. This study obtained significant evidence for the main effect of helpfulness ratio on user attitude toward the product or service, replicating the finding in experiment 1. Regardless of review types, high helpfulness ratio is more influential that low ratio of helpfulness votes in fostering user favorable attitude. The findings provide significant evidence for the generality of ratio effect.

2.5.3 Experiment 3

Experiment 3 had three main objectives. First, I tested another review helpfulness metric, i.e., vote magnitude. As vote ratio constant, I exploited whether the increase of helpfulness votes influenced user attitude. Second, I tested the boundaries of the proposed magnitude effect by examining both the positive and negative as well as attribute- and emotion-based reviews. Third, I tested the mechanism of the magnitude effect by measuring intermediator trust.

Given that experiment 1 and 2 have demonstrated that a review with high ratio is perceived as helpful and trustworthy, experiment 3 focused on a helpful review, the review with high helpfulness ratio, to test magnitude effect further. Besides, unlike Experiment 1 and 2, which used 50:6 and 52:4 to indicate the helpful review, here I set the high helpfulness ratio as 30:4 to further expand the generality of vote ratio effect independent of numerical values.

2.5.3.1 Method

I recruited 401 participants from Amazon Mechanical Turk. Each participant was compensated for her or his time with 1US \$. They were randomly assigned to one of eight conditions in a 2 (vote magnitude: large vs. small) × 2 (review valence: positive vs. negative) × 2 (review type: attribute-based vs. emotion-based) between-subjects factorial design.

The procedure is the same as that in Experiment 1. In the small magnitude condition, the participants were presented with a 30 "Yes" and 4 "No," whereas in the large magnitude condition, they faced 300 "Yes" and 40 "No." To tease out the magnitude effect, I make the ratio constant across small and large magnitude. As shown, the ratio of both 30:4 and 300:40 is 15:2. To enlarge the vote magnitude, I manipulate 340 total number of votes as large magnitude while 34 total number of votes as small magnitude. Then the large magnitude is 10 times the small magnitude. In this study, I assume this difference can distinguish the two levels of vote magnitude.

After reading the review, participants were asked to answer several questions. The participants completed a three-item measure of attitude toward the restaurant ($M=4.113$, $SD=2.096$) adapted from Rucker and Petty (2004) with the question, "What do you think of this restaurant?" The items were answered on a seven-point scale (1=Bad/Unfavorable/Dislike, 7=Good/Favorable/Like; $\alpha=0.990$). Next, participants completed the two-item measure of trust of the review ($M=4.363$, $SD=1.592$), e.g., "I trust this review," and items were answered on a seven-point scale (1=Strongly disagree, 7=Strongly agree; $\alpha=0.954$).

2.5.3.2 Results

Manipulation Check. For the manipulation check of review valence, participants were asked to indicate the extent to which they agree the review ranging from very negative to very positive on a seven-point scale (1=very negative, 7=very positive). The result of one-way ANOVA indicated that the manipulation of review valence was successful ($F(1,399)=1858.341$, $p<0.001$). The participants in the positive review condition considered the review as positive ($M=6.399$, $SD=0.935$), and the participants in the negative review condition considered the review as negative ($M=1.621$, $SD=1.264$). For the manipulation check of review type, one-way ANOVA yielded a significant main effect of content type ($F(1,399)=576.562$, $p<0.001$) in such a way that the participants in the attribute-based condition ($M=5.272$, $SD=1.592$) reported more than the participants in the emotion-based condition (M=1.764, SD=1.392). These findings suggest that the manipulations of both review valence and type were successful.

Attitude. I conducted three-way multivariate ANOVA and obtained a significant three-way interaction effect of helpfulness magnitude, review valence, and review type on attitude (F(1,393)=5.366, p=0.021<0.05). To clarify the three-way interaction effect, I first split the file by review valence. For negative review, I obtained the significant interaction effect of type and magnitude on attitude (F(1,393)=12.308, p<0.001), whereas the interaction effect was not salient for the positive review (F(1,393)=0.817, p=0.367). The file was further divided by review type. For the negative attribute-based review, I obtained the significant effect of magnitude on trust (F(1,393)=15.538, p<0.001), whereas the magnitude effect was not salient for the negative emotion-based review (F (1,393)=1.367, p=0.243). Finally, I conducted simple contrast for the negative attribute-based review and found that the review with large vote magnitude induced more unfavorable attitude (M=2.708 SD=1.375) than the review with small vote magnitude (M=1.718, SD=0.846; F(1,393)=19.864, p<.001), supporting H3a.

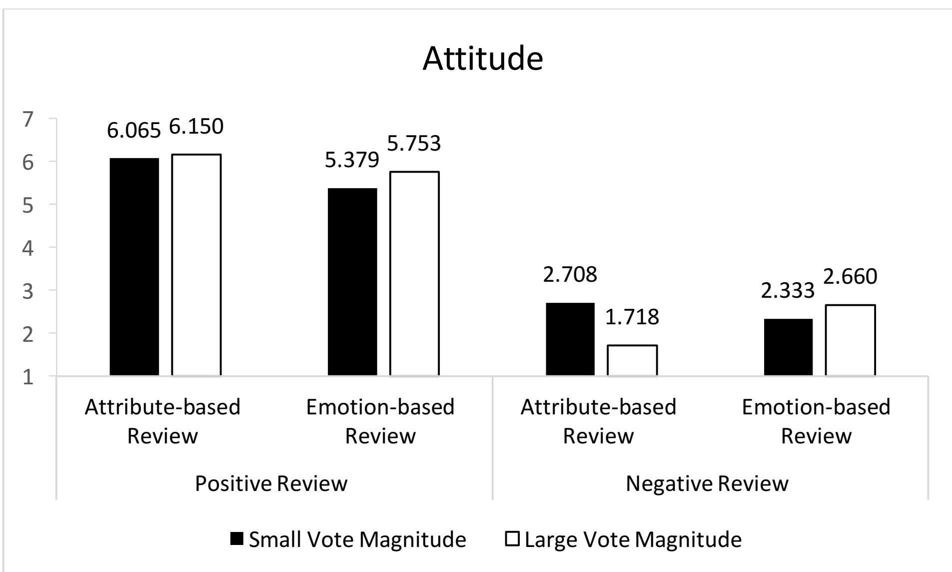


Figure 7. Impact of Vote Magnitude, Review Valence, and Review Type on Attitude

Trust. I obtained a significant three-way interaction effect of helpfulness magnitude, review valence, and type on trust (F(1,393)=4.552, p=0.034<0.05). To clarify the three-way interaction effect, I conducted simple contrast. There was no significant difference in users' trust for emotion-based negative review with small vote magnitude (M=3.917, SD=1.552) and the review with large vote magnitude (M=3.700, SD=1.542; F(1,393)=0.505, p=0.478). And there was either no significant difference in users' trust for emotion-based positive review with small vote magnitude (M=3.716, SD=1.718) and the review with large vote magnitude (M=4.150, SD=1.553; F(1,393)=2.091, p=0.149). And I didn't either obtain significant difference in users' trust for attribute-based positive review with small vote magnitude (M=4.990, SD=1.402) and the review with large vote magnitude (M=5.010, SD=1.541; F(1,393)=0.004, p=0.948). In contrast, there was a significant difference in users' trust for attribute-based negative review with small vote magnitude (M=4.354, SD=1.403) and the review with large vote magnitude (M=5.010, SD=1.330; F(1,393)=4.708, p=0.031<0.05), supporting H3b.

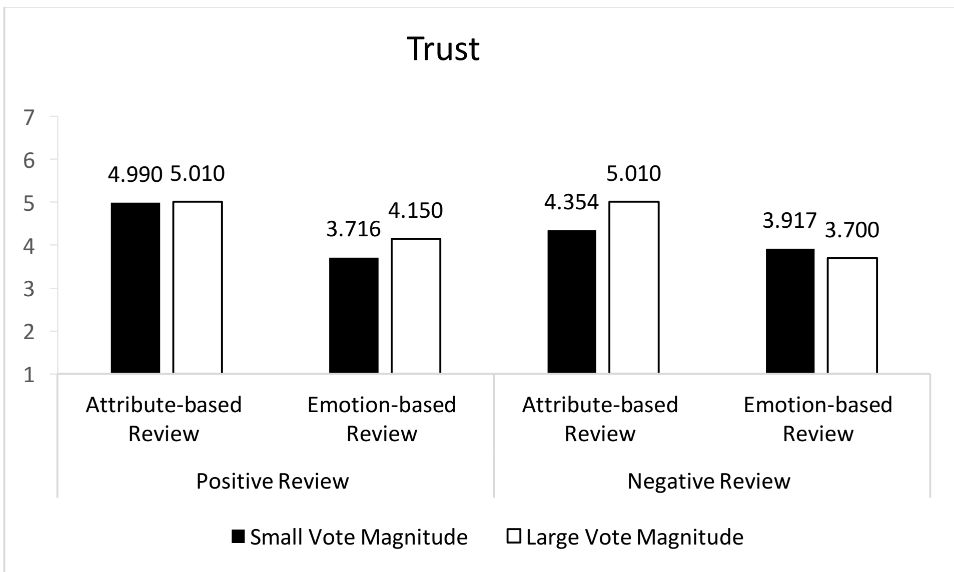


Figure 8. Interaction Impact of Vote Magnitude, Review Valence, and Review Type on Trust

Mediation Analysis. I applied Process Model 11 (Hayes 2013) and obtained the following results. The mediation effect of trust on the relationship between helpfulness magnitude and attitude (95% confidence interval: -0.4578 to -0.0099) is significantly moderated by review valence, which is moderated by review type. Furthermore, for negative and attribute-based reviews, trust significantly mediated the effect of helpfulness ratio on attitude (95% confidence interval: -0.3184 to -0.0055), supporting H3b.

3.5.3.3 Conclusion

My experimental results showed that compared with the review with small vote magnitude, the one with a large magnitude is more influential for the negative attribute-based review in strengthening user trust and decreasing the corresponding attitude. Unlike vote ratio as a direct trust cue, the results suggest that the effect of magnitude depends on the valence and type of the review. When users are motivated to elaborate the review, they will use magnitude as reference. Given that the negative review is perceived as more risk-taking as reference for product evaluation, users need additional information to justify the validity of the negative comments in the review. Likewise, the attribute-based review elicits user interest to obtain complete validation.

2.6 Discussion

Opinion evaluation takes place in widely used participatory platforms, ranging from social media sites to third-party review websites. For instance, Facebook provide a “Like” button to allow users to express their endorsement for posts of others. Besides, YouTube offers users a mechanism to either “thumb-up” or “thumb-down” a video. Perhaps opinion evaluation is most apparently exemplified by the voting system on third-party review sites (e.g., Amazon.com, Dine.com). Reasoning on the review voting is fundamentally different from reasoning in the review itself. For example, instead of asking “What did you think of the product?” the voting system asks, “What did you think of the other user’s opinion on the product?” The vote metrics that are either presented by ratio or magnitude provide subsequent users a direct indicator of the extent to which prior users perceived the review as helpful. The current study investigated how online users respond to prior users’ helpfulness votes in e-commerce context.

My results indicate the selection of cues among different heuristic cues. That is, the processing of helpfulness votes follows a sequence, i.e., ratio will be utilized first and magnitude will be used later when users have the motivation and ability to internalize it as reference in decision making. Concretely speaking, experiment 1 verifies the effect of vote ratio across positive and negative reviews. A positive review with a high helpfulness ratio is more influential than that with a low ratio of helpfulness votes in fostering the favorable attitude of customers. By contrast, a negative review with a high helpfulness ratio is influential in driving unfavorable attitude. The result also indicates the mechanism of the ratio effect, namely, trust on the review. Experiment 2 further validates the generality of ratio effect across attribute-based and emotion-based reviews. Experiment 3 examines the magnitude effect. I find that unlike vote ratio, which is a basic and consistent trust driver, the effectiveness of magnitude depends on the valence and type of the review. A negative review, which contains sensitive and attention-tracking information, is likely to motivate users to consider magnitude as reference. Likewise, attribute-based review, which contains informative and objective information, enables users to take advantage of magnitude as reference.

2.6.1 Theoretical Implications

I focus on the link between helpfulness votes and user evaluations. I suppose that helpfulness vote ratio can directly function as trust cues, whereas helpfulness magnitude partially acts as trust cues regardless of review valence and type. Compared with prior literature, which only focuses on the effects of online reviews (Duan et al. 2009; Ludwig et al. 2013), the determinants of review helpfulness votes (Mudambi & Schuff 2010; Yin et al. 2014; Cao et al. 2011), and social influence (Kuan et al. 2014; Lee et al. 2015; Muchnik et al. 2013), the present study contributes to literature on online product reviews and social influence in several ways.

First, the findings add to extant research that examines the commercial value of online product reviews. Online product reviews, which refer to informational communications among users concerning evaluations of goods and services, play an increasingly important role in electronic commerce in various aspects, such as informing potential users of product knowledge (Martin & Lueg 2013), reducing uncertainty of product quality (Senecal & Nantel 2004), and increasing product sales (Chevalier & Mayzlin 2006). My findings identify another cue, i.e., helpfulness votes with commercial value. My results indicate the selection of cues among different heuristic cues. Hence, the processing of helpfulness votes follows a sequence, i.e., ratio will be utilized first and magnitude will be used later when users have the motivation and ability to internalize it as reference in decision making.

Most studies on review helpfulness aimed to identify the determinants of review helpfulness (e.g., Mudambi & Schuff 2010; Yin et al. 2014; Cao et al. 2011); only a few studies directly examine the metrics of helpfulness votes and their consequences. For instance, Purnawirawan et al. (2012) demonstrated the association between perceived review helpfulness and user attitude and intention, which indirectly implies the importance of helpfulness of reviews. However, perceived review helpfulness is not an appropriate approach to illustrate the numeric characteristics of review helpfulness in online review platforms. Chen et al. (2008) directly figured out the higher influence of reviews with a high proportion of helpful votes for less popular books. However, both studies do not provide a complete and direct examination of helpfulness information itself. Thus, I add to the current helpfulness literature by providing a direct depiction of review helpfulness and justify its commercial value.

Finally, I extend the current research on social influence theory in two aspects. In terms of manipulating social influence, I employ an accurate factor, i.e., ratio of helpfulness votes to illustrate the valence of social influence, rather than adopting positive or negative opinions of others to reflect the valence of social influence (Graziano et al. 1993) or simply viewing up-votes as a positive influence and down-votes as a negative influence (Muchnik et al. 2013). Regarding the consequences of social influence, existing studies examine herding effect on voting behavior (e.g., Muchnik et al. 2013), whereas I move one step from primitive voting behavior of subsequent users to a more commercial purchasing behavior.

2.6.2 Practical Implications

I believe that my work has a number of important implications for multiple aspects that range from voting system providers who are interested in pursuing information techniques on third party review platforms to the general participatory sites, which highlight the user-generated content sharing and dissemination, and finally to the online retailers who highly value awareness and reputation of products.

Given its simplicity, the helpfulness voting system is crucial on platforms, such as Yelp (e.g., number of "useful," "cool," and "funny" votes), Dine (i.e., number of both helpful and unhelpful votes), and Amazon (e.g., percentage of people who found the review helpful). The results suggest that the helpfulness magnitude takes forces on condition, whereas helpfulness ratio is an unconditional cue to elicit user trust on the reviews of others and initialize them as reference for decision making. Therefore, the voting system providers can expose helpfulness ratio or simply demonstrate the number of helpful votes.

Consumers who view webpages likely see information about others who voluntarily support the brand or a comment by clicking social buttons, such as the "like," "share," and "comment" buttons. With the proliferation of different types of social buttons across social media websites, a number of predefined user activities (e.g., voting, recommending, and sharing, tweeting, liking) are taking effect on social media platforms. These buttons often show a counter of the number of times a post has been shared, liked, or recommended, such as x likes, y dislikes, x shares, x tweets, which indicate user-generated numbers. These platforms simply display the number of one-sided votes without a clear presentation of a vote ratio. Using the current results, these platforms can display the number of "views" on a post, which may provide a proxy for the ratio and magnitude between "up-vote" and "down-vote." Users may perceive these user-generated numbers considering the unconditional ratio rather than the conditional magnitude perspective.

Increases of "up-vote" and the presence of "down-vote" are crucial given the importance of attention and awareness for the success of online retailers, especially the middle and small ventures. To achieve higher reputation and aim for the first-page display in user searching on third-party platform (e.g., Taobao.com), some retailers will buy fake "up-vote" and eliminate "down-vote." However, my results show that magnitude is influential when opinions are negative and in an attribute-based style. Thus, the investment on forum irrigation people may not have high return of investment.

2.6.3 Limitations and Future Research

This study can be further strengthened in several ways. First, the effect of the voting system is not fully explored in this research. To demonstrate the valence and strength of normative influence endowed by helpfulness votes, I examine the changes in the numbers of up-votes and down-votes on a review. However, various voting systems only provide a "thumb-up" or "like" option to allow users to participate in review evaluation. Thus, the publicity of prior user votes is restricted to positive influence, which prevents subsequent users from performing a rational trade-off between up-votes and down-votes. Whether such a voting system has commercial value, as the discussed one in my research and if one-sided votes influence response of other users remains an interesting research question. Another eye-tracking cue, i.e., "view" counter, is often displayed beside the counter of up-votes on a review. The difference between the number of views and votes may have varying user interpretations. For example, the difference can be perceived as the number of users who down-vote the review, which possibly engenders similar findings of my research. However, users may have low motivation to conduct voting action or have neutral attitude toward the review. Thus, separating the two possibilities is empirically important to answer.

Second, I analyze the effects of helpfulness votes at the individual level. I conduct experiments to explore ways individual perceptions of helpfulness votes influence their trust of the review, downstream product evaluations, and purchase intention. Examining how users process and react to helpfulness votes in different ratios and magnitudes is important to understand the effectiveness of the review voting system as a marketing tool. However, I did not model the effects of helpfulness votes on product sales at an aggregate level. Quantifying the relationship between changes in helpfulness votes and product sales through time is important to provide in-depth understanding of the dynamics of vote formation and its corresponding commercial value. For the experiments, I have two studies using MTurk samples and one study using student sample. In other words, I use both MTurk and Student samples to test the hypotheses with purpose of justifying that my identified phenomenon not only exists among workers but also students, both of whom may have online shopping and reviewing experience. In the future, I will replicate the three studies using MTurk samples and then use the student sample to cross-validate the results.

Third, I theorize that vote magnitude is effective on the condition of both motivation and ability. That's why I propose a three-way interaction rather than two two-way interactions. Maybe in the future I will explore the two-way interactions between vote magnitude and review valence, and between vote magnitude and review type. Regarding the vote ratio, I hypothesize it is a unconditional cue, regardless of review valence and type. In order to clearly tease out the main effect, I propose two hypotheses separately arguing the ratio effect independent of valence and type. Maybe I will test the three-way interaction in the future.

Finally, this study focuses on the voting system in the third-party review platform, that is, the evaluation on a review about a product or service. However, the voting system is widely applied to online settings, ranging from likes of a post on Facebook, to thumb-ups of a microblog on Twitter, and to up-votes of a comment of the article on The New York Times. The contextual moderators in my findings should be explored further given the possible distinctions of contexts.

Chapter 3

The Impacts of User-Generated Photos on Consumer Information Processing and Decision Outcomes

3.1 Introduction

E-Word of Mouth (e-WOM) spreads via social media and comprised of user-generated content (UGC) pervade economies, society, and organizations. E-WOM comes in many forms, including picture, text, and video. While text e-WOM has existed since the first e-mail, the growth of photo and video sharing has been striking on platforms such as Facebook, Instagram and Flickr. For example, Facebook users upload more than 350 million photos each day[3]. More interestingly, user-generated photos are significantly different from pictures in the traditional marketing communications such as advertisements and formal product descriptions. Accordingly, I identify four key characteristics of user-generated photos: 1) *Ubiquity*—a huge amount of photos have diffused into every online review platforms such as Yelp.com and Openrice.com; 2) *Varied Quality*—most of the photos are taken and posted by users, not by professionals, thus of varied visual quality; 3) *Varied Presentation*—these photos are presented with text in different layouts—e.g., sometimes intertwined with the main text vs. being separated from the main text, and sequences—e.g., sometimes followed by the main text .vs following the main text; *4) Varied Density* —these photos are presented with text in different topical density—e.g., sometimes many certain topic-based photos matching with the text vs. a few photos corresponding to the text.

The pervasive use of e-WOM to share photos has significant implications for business, because customers make decisions based on both text and photo. In particular, in the context of e-WOM, diners at restaurants take photos of food and facilities and post them to ratings platforms such as Yelp.com or Ricebowl.com. Often, potential diners view photo more trustworthy than pictures in advertisements, because they are seen as authentic representations of what dishes look like when they appear on the table (Goh et al. 2013). That potential customers view photo as trustworthy, creates a pressing need research that offers a richer understanding of how photos influence consumers' decision-making in the context of e-WOM.

[3] http://www.businessinsider.com/facebook-350-million-photos-each-day-2013-9.

Existing research on e-WOM has been mainly focusing on the impacts of numeric ratings & textual content in online customer reviews on consumer information processing and related business outcomes (e.g., Chevalier & Mayzlin 2006; Dellarocas et al. 2007), however, insufficient research ever pays attention to the photo e-WOM. Even though there are studies exploiting the role of pictorial stimuli in consumer information processing and decision making, for example, Kim et al. (2008) emphasizing the superiority of photo in attitude formation while Jiang & Benbasat (2007b) identifying the superiority of the combination of both photo and text based on the cue summation theory, little is known about the interaction between photo and text in presentation, quantity and topics. Furthermore, these studies mainly focus on the standard product/service descriptions and advertisements produced by business rather than e-WOM developed by customers. Then how about the impact of photos generated by customers on others' information elaboration and downstream decision-making?

To answer the above questions, I propose that the addition of pictures in e-WOM may not always create instrumental effects, depending on the photo presentation (e.g., layout, sequence) and density. That is because the supplement of photos may create information interference brought by not suitable framing and presentation of photo and text. Then I assert that the additional photos sometimes inhibit the understanding of the e-WOM, depending on the presentation and density of photo relative to text. My research mainly has three objectives: (1) shed light on the integration impacts of text and photo on consumer information processing and decision making outcomes on ratings platforms; (2) test the mechanism behind the presentational impact of photo using an experimental study; (3) employ field aggregate-market data to reveal the effectiveness of photo e-WOM on organization performance; (4) provide guidance to managers in leveraging photo to improve the effectiveness of e-WOM.

In terms of photo presentation, I propose three representative hypotheses on the presentational integration of photo & text on information processing and decision outcomes, based on the cue summation theory of multiple-channel communications (Severin 1967). In particular, I examine interaction effects of the relative layout (the alternate layout vs. the separate layout) and the relative sequence (text first vs. photo first) on customers' perceived diagnosticity, pleasantness of e-WOM and their attitude towards the recommended product/service in e-WOM. I conduct an experimental study to provide a robust test of my hypotheses. Overall, I mainly find that 1) for perceived diagnosticity, separate layout is better than the alternate layout, especially when the picture is displayed first than the text first; 2) for pleasantness, alternate layout is better than separate layout, regardless of the sequence of text and photo; 3) text first presentation does better than photo first because text acts as a guide in information understanding.

Concerning photo density, my research aims to examine the interaction effects of overall photo density (i.e., average number of photos per review), inside &outside & food & drink & menu photo densities (i.e., average number of in inside/outside / food /drink / menu photos per review) and restaurant niche width based on aggregate-restaurant data. To clarify my interaction effects, I use stepwise regression with three blocks of variables: control effects, linear effects, and interaction effects. Overall, I find that that (1) sharing more photos especially outside photos in a review hurts restaurant reputation while sharing photos more on food, drink and menu of a restaurant increases its reputation; and (2) for the restaurant as a generalist occupying multiple cuisines, the more photos shared in a review, the better a restaurant reputation will be; (3) by contrast, for the restaurant as a specialist occupying few cuisines, the more photos shared on food and drink in a review, the better a restaurant reputation will be.

In summary, the experimental study examines the interaction between photo and text from presentational perspective while the field study exploits their interaction from density perspective. Simply speaking, I conduct the lab-based experiment to accurately figure out which kind of layout and sequence of photo relative to text is better in facilitating the understanding and persuasiveness of a review. That is, the experimental study is review-level analysis. In order to have a more general and comprehensive exploration of the interaction between photo and text, I conduct aggregate organization-level analysis in the field study. Given that online customer reviews are generated to recommend or assess the quality of items provided by the organization, it is also necessary to know the proximal match between the number of photos and the number of reviews for an organization. In a word, the experimental study uncovers the interaction between photo and text in a micro-way—i.e., individual review level while the field study provides a macro-way to see their interaction by aggregating reviews into organization level.

These findings make important contributions. First, current cue-summation theory mainly focuses on two factors—i.e., number of cues and relevancy of cues involved in cue summation (Jiang & Benbasat 2007a; Dimoka et al. 2012). My research extends cue-summation theory (Severin 1967) by giving a look at the efficiency and ease of cue summation. Concretely speaking, I focus on between-channel interactions in the cue summation process and investigate how presentational layout and sequence of photo relative to text affect the efficiency and ease of cue summation. And I propose that the layout could influence the efficiency, i.e., compared to the alternate layout, the separate layout with one time channel-switching increases the efficiency of cue summation. Besides, the sequence could affect the ease of cue summation, i.e., the text-first sequence followed by photos can make the summation of different channel cues easy. Because text can function as guidance in reading and alleviate the possible noise brought by photos. Second, my research also extends e-WOM literature (e.g., Godes & Mayzlin 2004) by incorporating the examination of photo e-WOM and contributes to the multimedia communication literature (e.g., Kim & Lennon 2008; Dimoka et al. 2012) by providing insights into the effectiveness of photos in the e-WOM context. Third, the field study empirically revealed that photos did not necessarily generate more positive impression, dependent on the content of photos (e.g., food-relevant or environment-relevant) and restaurant niche width, which extends the prior literature on the photo-superiority effect (Peterson & McGee 1974; Edell & Staelin 1983; Kim & Lennon 2008). Fourth, my findings add to the photo literature which examines visual quality (e.g., athletics, size, complexity) (Deng 2010; Cyr et al. 2009) by considering the content of photo—i.e., food, drink, inside, outside, menu. Practically, my findings offer some actionable implications for managers of online review websites (e.g., Yelp.com) in critical information screening and selection, online retailers in providing guidelines for customer photo sharing and managers of restaurants in selection of cues (e.g., food or decorations) to expend marketing budgets for the maximum return of investment (ROI).

3.2 Literature Review

3.2.1 Online Rating System

Online rating system carries e-WOM on products or services to establish the reputation of the organization (Ba & Pavlou 2002). Online review websites mostly deploy the system to provide a platform where buyers share and exchange reviews on products or services (Fang 2012). As a representation of buyers' overall evaluation of the value of organizations and their products/services, consumer value ratings (e.g., 5star rating) are increasingly declared in public forums through online reviews that are usually posted on review websites (e.g., Yelp.com, CNET.com, Amazon.com) (Zeithaml 1988, p. 14).

As suggests by Luca (2011), a one-star increase in Yelp ratings leads to a 5%–9% increase in revenue of the organization. Concerning the impacts of consumer ratings on purchase behavior or sales revenues, there are three frequently studied rating metrics, which are, volume (Liu 2006), valence (Chevalier & Mayzlin 2006; Dellarocas et al. 2007) and variance (Sun 2012). However, except for the research on the impacts of rating and its metrics, little is known about the determinants of consumer value rating, which is an important index for organizational reputation (Luca 2011; Kovács et al, 2013).

3.2.2 Photo e-WOM

Electronic Word-of-mouth (e-WOM) is defined as oral, customer-to-customer online communication on a brand or product, a service or a provider (Arndt 1967). E-WOM has been shown to positively influence buyers' purchase behaviors (Dellarocas et al. 2007; Chevalier & Mayzlin 2006). However, previous studies have mostly focused on e-WOM associated with text reviews and numeric ratings (Dellarocas et al. 2007). The impact of photo e-WOM generated by buyers on reputation platforms has not been well exploited.

Photos and texts are two crucial carriers of information, the former acting as the visual communication while the latter as the verbal communication, but they are processed differently (Jiang & Benbasat 2007a). Compared to text, pictures are in principle more attention-getting, easier to process and understand in a holistic manner, and generate stronger feelings—i.e., the picture superiority effect (Peterson & McGee 1974). Research on photos or visual communication can be classified into three streams: (1) examine the differences between verbal and visual communication and identify the superiority of the combination impacts of verbal and visual stimuli (Peterson & McGee 1974; Edell & Staelin 1983; Kim & Lennon 2008; Jiang & Benbasat 2007a); (2) exploit photographical features, such as vividness, complexity, order, human image (Jiang & Benbasat 2007b; Deng 2010; Cyr et al. 2009); (3) explore the power of the aesthetics quality of photos(Aydin et al. 2015). However, little is known about the interaction between photo and text in quantity and content. Moreover, insufficient studies ever examine the content of photos. In other words, not only the number of photos in a review matter, but also the content of photos in a review plays a role.

While e-WOM research has simply focused on text and ratings, a complementary stream of research suggests that pictorial stimuli have an equally powerful effect on consumer information processing and decision making. Findings suggest that pictorial stimuli have a stronger impact on attitude formation than text (Kim & Lennon 2008), because pictures stimulate the imagination and are more likely to elicit enjoyment from seeing the actual consumption and therefore trigger a desire to consume a product or service. In other words, compared to text, pictures are in principle more attention-getting, easier to process and understand in a holistic manner, and generating stronger feelings (Peterson and McGee 1974). Recent studies argue that combining picture and text (sometimes referred to as verbal) increases information comprehension (Jiang & Benbasat 2007a; Dimoka et al. 2012).

In summary, existing research mainly examine two forms of e-WOM—i.e., ratings and text, however, the effectiveness of photos is not well investigated. Further, despite some studies identifying differences in pictorial stimuli and textual stimuli in information comprehension, more insights into the integration of picture and text are needed. Furthermore, rather than photos generated by consumers, most studies of photos focus on professionally produced product/service. Given the identified research gaps, I aim to investigate the integration impacts of photo and text on consumer decision-making and evaluation outcomes from the presentational perspective—e.g., the relative layout, sequence and density of photo.

3.2.3 Organizational Niche Width

An organization can shape its identity by positioning itself into existing specific categories known as “niche width” (Kovács et al, 2013). An organization with a small niche width serves a small amount of the market, catering to specialized demands, which is often regarded as a specialist. By contrast, an organization with a big niche width serves a large amount of the market, catering to various demands, which is often regarded as a generalist. Customers often rely on categorization to identify and interpret the products or services provided by the organization. For example, restaurants spanning multiple cuisines (e.g., serving both Chinese dishes and Japan sushi) often display features unlike those of any of the single-cuisine restaurants (e.g., only serving Hamburger) (Kovács et al, 2013). And audiences perceive organization spanning multiple categories as a signal of ambiguous identity and brand position (Hsu et al. 2009). Thus, organizational niche width is noteworthy in the examination of the antecedents of the organizational reputation.

3.3 Experimental Study

3.3.1 Theoretical Background and Hypotheses

In this study, I draw mainly on the cue summation theory of multiple-channel communications (Severin 1967) to develop my research model and hypotheses. Here, a channel is equivalent to a modality of communication (Grifoni 2009), such as a text channel that consists of textual cues and a picture channel of pictorial cues. The cue summation theory posits that in principle the communication effectiveness depends on the summation of cues from different channels—the more channels involved, the better the communication. This superiority of multi-channel communication has found the support from quite a number of IS studies (e.g., Jiang & Benbasat 2007a & 2007b).

In the current study, I focus on such between-channel interactions in the cue summation process. I point out three premises before developing my research model and hypotheses in this section. First, I follow the premise that compared to text, pictures are in principle more attention-getting, easier to process and understand in a holistic manner, and generating stronger feelings—i.e., the picture superiority effect (Peterson & McGee 1974). Given that pictures facilitate both cognitive understanding and affective feelings, I examine two types of information processing outcomes in the current study—the cognitive outcome (diagnosticity) and the affective outcome (pleasantness). Perceived diagnosticity is defined as the extent to which users believe the overall review is helpful to evaluate products or services (Kempf & Smith 1998). Pleasantness is defined as a positively valence of emotion (Deng & Poole 2010). Second, I focus on e-WOM that consists of multiple paragraphs and pictures—e.g., paragraphs and pictures about the decoration, the taste of food, the service, etc. of a restaurant. Third, I identify two major presentational features—i.e., relative layout and sequence of pictures. The alternate layout groups text and photo by topic while the separate layout puts all the text before the pool of photos. In terms of relative sequence, I focus on two levels—i.e., text first with text displayed before the picture while photo first with photo presented ahead of text. And I propose that the presentation and sequence of photo relative to text in a review could influence the efficiency and ease of cue-summation according to the cue-summation theory of multi-channel communication. Regarding other aspects of photos, such as color and size which can reflect the visual aesthetics, I agree that they can affect the interpretation of photos. Given that my focus is on the interaction between photo and text, I think the interactional factors such as layout and sequence of photo relative to text more carter to my research objectives. Finally, I focus on positive e-WOM in my current study though my model can be readily adapted to incorporate the negative e-WOM. Figure 9 depicts the research model.

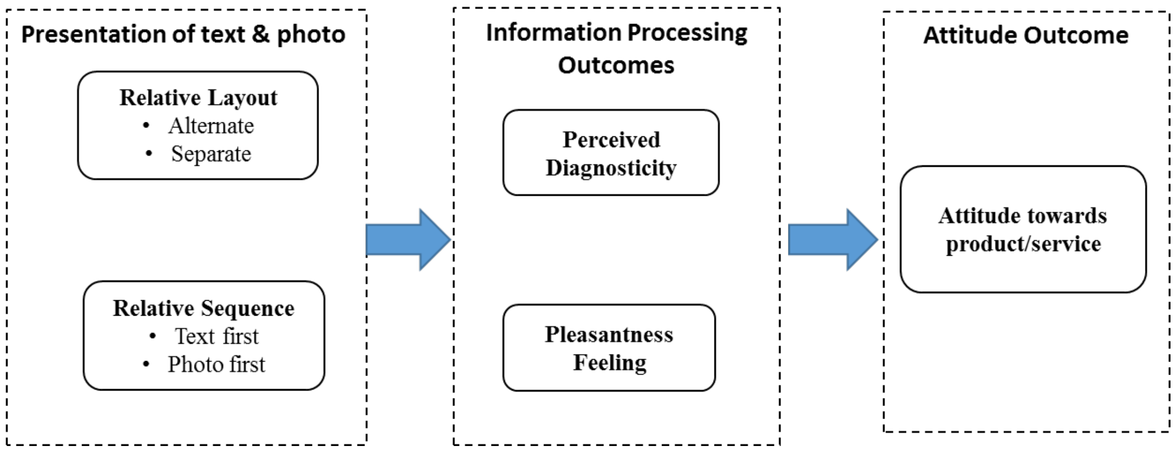


Figure 9. Model of Presentation of Photo e-WOM

3.3.1.1 Sequence Impact

I posit that that perceived diagnosticity of an e-WOM is mainly influenced by the ambiguity of the e-WOM as a result of the summation of both textual and pictorial cues. This is because perceived diagnosticity depends on whether consumers can have clear information to support their final decisions (Kempf & Smith 1998). While pictures are easier to process and understand, they may contain noise for a consumer without the right framing or guidance of the textual cues, thus decrease the efficiency of cue-summation. Thus, I posit that the presentational sequence of contents between text and picture channels affects cue summation efficiency: in text first presentation, text acts as guidance for processing picture(s) and helps develop a clear understanding of the e-WOM while in photo first presentation, no guidance of text to process picture(s), which hurts the elaboration of the e-WOM. Thus,

H1: E-WOM in text-first presentation will be perceived of higher diagnosticity and lead to more favorable attitude towards the product/service than the one in photo-first presentation.

3.3.1.2 Layout Impact

Given that in text first presentation, text acts as guidance for processing picture(s) and helps develop a clear understanding of the e-WOM. Thus, texts and pictures can work together to facilitate information processing and comprehension, regardless of the relative layout. However, when photo is firstly presented, no guidance of text to process picture(s), which hurts the elaboration of the e-WOM. Thus, there can be nonetheless costs brought by the interference between channels to alleviate the efficiency of the overall cue summation.

I conjecture that in the photo-first presentation, the relative layout of text and photo interact to affect perceived diagnosticity of the e-WOM by influencing the efficiency of cue summation. In the separate layout, all the texts and pictures are clustered in two separate pools, which suggests one time of channel-switching in cue summation process. Especially when all the texts are presented in the beginning, which helps generate a comprehensive map of the product/service before interpreting pictures, customers are more likely to be able to recall what they have read with the reminder of subsequent pictures and thus solidify their understanding. However, in the alternate layout, the textual cues will interfere more frequently with the pictorial cues because consumers need to switch between the text and the picture channels from time to time, which increase costs of channel-switching and decrease the efficiency of cue summation. Thus,

H2: In the text-first presentation, the relative layout is not salient; In the photo-first condition, the e-WOM in the separate layout will yield higher diagnosticity and more favorable attitude towards the product/service than the one in the alternate layout.

In addition, I conjecture that the relative layout of text and photo also interact to affect pleasantness feeling by influencing the ease of cue-summation. Pictorial cues via imagination are likely to elicit enjoyment from seemingly actual consumption and trigger desire for a product or service (Kim & Lennon 2008). I posit that the between-channel interference will also influence the outcome—that is, while pictures are generally more attention-getting and producing enjoyment, the relatively layout may affect the level of pleasantness generated from cue summation.

In the alternate layout, consumers can easily cross-reference the textual cues when processing the pictorial cues and easily establish the connection between text and picture in the cue summation process. As a result, the intertwined textual cues strengthen the imagination of seemingly actual consumption elicited by pictures that vividly represent a product or service. In contrast, the separate layout 'sum up' all the pictures in one pool, which creates sort of isolation between text and picture channels, thus reducing the ease of cue summation and finally inhibiting the smooth development of imagination. Thus,

H3: *Regardless of the relative sequence, e-WOM in the alternate layout will bring more pleasantness feeling and lead to more favorable attitude towards the product/service than the one in the separate layout*.

3.3.2 Method

I tested my hypotheses in an experimental study. This experiment had three main objectives. First, it tested the main effect of layout and sequence on perceived diagnosticity, pleasantness feeling and attitude towards the product/service; Second, it tested the interaction effect of layout and sequence on these outcomes; Finally, I also explored the mechanism behind the hypothesized layout and sequence effects.

3.3.2.1 Design

I recruited 162 participants (95 female, Mean (age) =22) from Amazon Mechanical Turk. For the detailed sample characteristics (e.g., education, income), please refer to Appendix E1. They were randomly assigned to one of 4 conditions in a 2 (Layout: separate vs. alternate) × 2 (Sequence: photo first vs. text first) full factorial between-subjects design. Each participant was compensated for her or his time with 1US $.

3.3.2.2 Procedures and Measures

At the beginning of the experiment, participants first answer two questions on their liking of Japanese food and eating experience in Japanese restaurants. And then they read the instructions and learn that the study is about evaluating a Japanese restaurant based on online customer reviews. The participants will be asked to role-play in the following situation: "Imagine that you want to select a good restaurant for dinner. After searching online, you find the following review about a restaurant". Subsequently, they were randomly exposed to one of the 4 manipulations as depicted in Appendix E3. After reading the review, participants will answer questions about the key variables in my model. Existing scales are adapted for my study. All scales are shown in Appendix E2.

Table 7. Descriptive Statistics and Correlations

Variables	Reliability	Mean	Std.Dev	1	2	3
1. Attitude	0.951	5.370	1.860	1		
2. Diagnosticity	0.593	5.317	1.431	.670**	1	
3. Pleasantness	0.936	5.681	1.193	.076	.523**	1

3.3.3 Results

Manipulation check. For the manipulation check of layout, participants were asked to indicate to what extent they agree the photos and text ranging from irrelevant to relevant on a 7-point scale (1=irrelevant, 7=relevant). The one-way ANOVA result indicated that the manipulation of layout was successful ($F(1, 160)=4.794$, $p<.05$): the participants in the separate condition considered the review more relevant (Mean=5.825) than those in the alternate condition (Mean=5.3415).

Perceived diagnosticity. I conducted a two-way ANOVA and got significant main effect of layout ($F(1, 158)=17.5$, $p=.000$) and sequence ($F(1, 158)=28.6$, $p=.000$) on perceived diagnosticity. Besides, I also obtained significant interaction effects between layout and sequence on diagnosticity ($F(1, 158)=37.5$, $p=.000$). As shown in Figure 10, when UGT first, layout didn't significantly change their diagnosticity ($F(1,158)=1.982$, $p=.161$). In contrast, when PHOTO first, separate layout significantly increased participants' perceived diagnosticity (Mean=5.75) than those in the alternate layout condition ($F(1,158)=50.588$, $p=.000$; Mean=3.85).

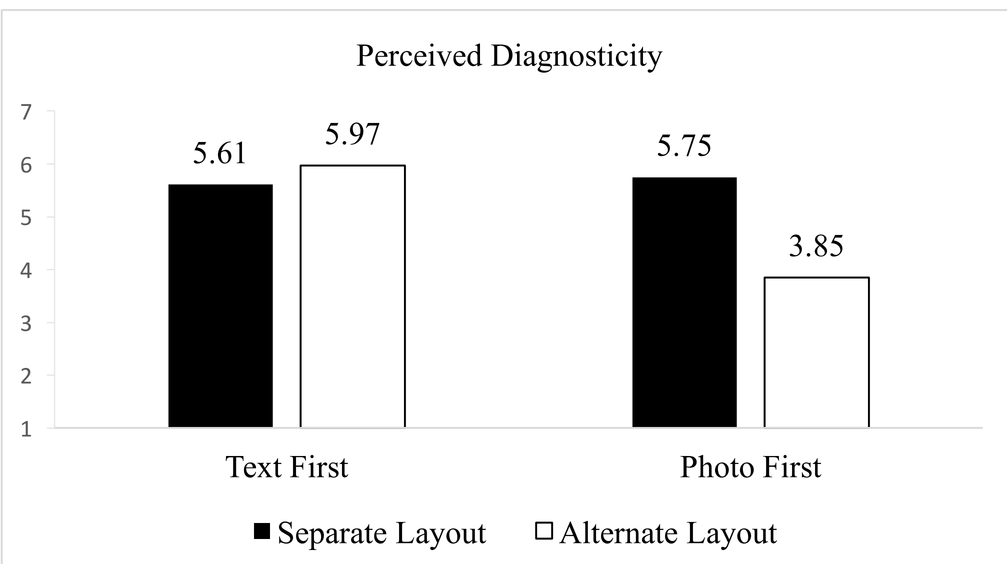


Figure 10. Interaction Impact of Sequence and Layout on Diagnosticity

Pleasantness. I only got significant main effect of layout (F(1, 158)=5.24, p=.023). As shown in Figure 11, regardless of sequence, alternate layout significantly increased participants' pleasantness than those in the separate layout condition.

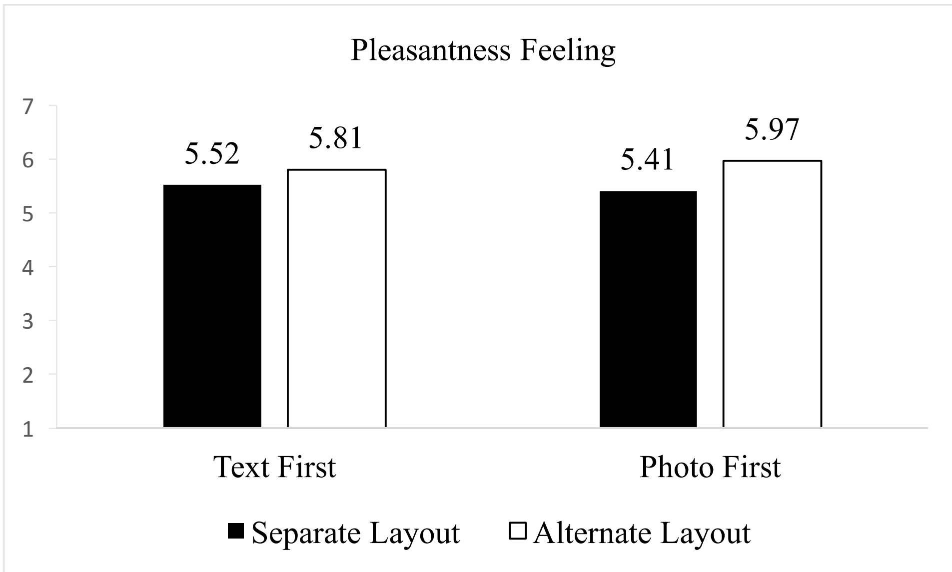


Figure 11. Interaction Impact of Sequence and Layout on Pleasantness

Attitude towards Restaurant. I got significant main effects of layout (F(1, 158)=68.692, p=.000) and sequence (F(1, 158)=96.092, p=.000) and also obtained significant interaction effects between layout and sequence (F(1, 158)=92.683, p=.000) on attitude. As shown in Figure 12, when UGT first, layout didn't have a significant impact (F(1,158)=.943, p=.333). In contrast, when photo first, separate layout significantly increased participants' attitude (Mean=6.08) than those in the alternate layout (F(1,158)=152.924, p=.000; Mean=2.79).

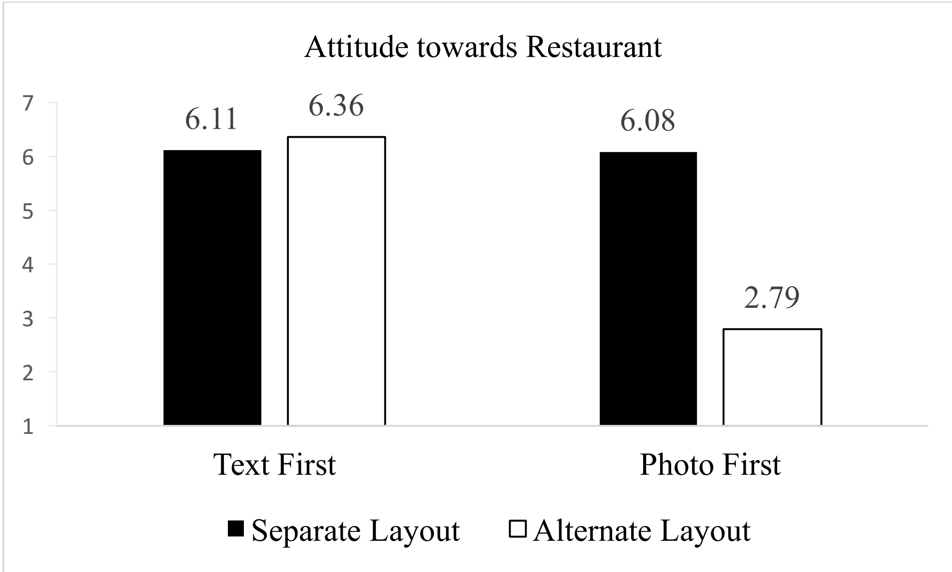


Figure 12. Interaction Impact of Sequence and Layout on Attitude

Mediation Analysis. I applied Process Model 8 (Hayes 2013) and got results as follows. The mediation impact of diagnosticity on the relationship between layout and attitude was significantly moderated by sequence (95% CI: -1.472 to -.553). That is, when text first, diagnosticity didn't significantly mediate the layout effect (95% CI: -.386 to .086) while when photo first, diagnosticity significantly mediated the layout effect on attitude (95% CI: .452 to 1.270).

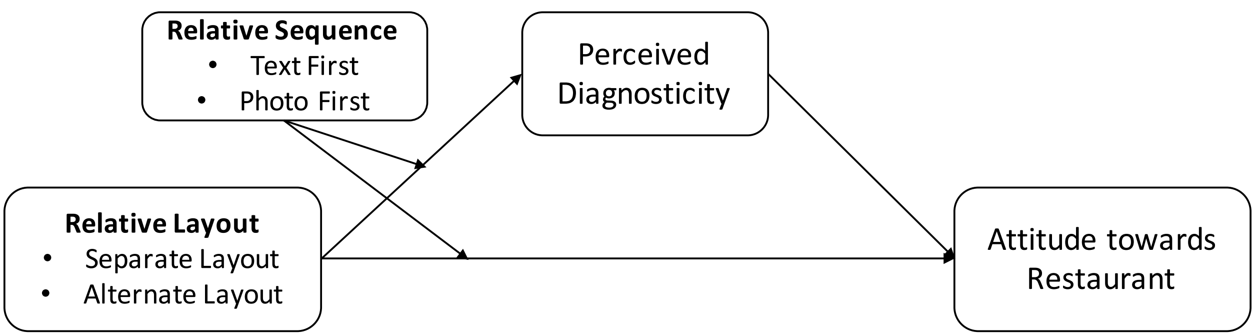


Figure 13. Moderated Mediation Analysis Model

3.4 Field Study

Different from text e-WOM, which needs cognitive resources for processing and prior knowledge to interpret (Edell & Staelin 1983), photos are able to vividly and intuitively communicate the performance of a restaurant and easily to arouse consumption-relevant imagination (Jiang & Benbasat 2007a). Then it arises a question, will the photo sharing accompanied with a textual review influence the overall reputation of a restaurant? If so, how many photos to match with the text will benefit the reputation? Further, what photos to share will improve or hurt the reputation, more food-relevant photos or more environment-relevant photos? Given that restaurants may endow different weights on the components of their reputation, that is, some specialized restaurants focus on certain typical food or drink while other generalized restaurants care about the whole development (Kovács et al 2013), then will the influence of photo e-WOM on the restaurant reputation varies by its cuisine spanning?

Concerning photo density, my research aims to examine the interaction effects of overall photo density (i.e., average number of photos per review), inside &outside & food & drink & menu photo densities (i.e., average number of in inside/outside / food /drink / menu photos per review) and restaurant niche width based on aggregate-restaurant data.

3.4.1 Data

3.4.1.1 Data Collection

The research context is a popular online review website Yelp.com founded in October 2004, which covers a broad range of 22 product and service categories such as restaurants, shopping, beauty & spas, public services, etc. I used Yelp Challenge Data Set (https://www.yelp.com/dataset_challenge) released in June 2016. Based on the dataset, I built up my own research sample. First, I selected restaurants as the research object because restaurant is typical experience goods whose quality can't be thoroughly inspected before purchasing. Then reviews or recommendations from other buyers play an important role in the other consumers' assessment of the product/service. Second, I selected Las Vegas as the target city which is most popular with the largest number of reviews in the dataset. Thus, the sample has 96,588 photos in 1,241,046 reviews from 15,517 restaurants from Jan 2004 to June 2016, within which 13,995 photos are on the inside, 12,865 photos on the outside, 15,329 photos on the food, 14,764 photos on the drink, 15,035 photos on the menu and 24,600 photos belong to none of the categories.

3.4.1.3 Variables and Measurement

The dependent variable is restaurant reputation, which is operationalized as the average consumer value rating of a restaurant. The independent variables of interests are the photo density, inside photo density, outside photo density, food photo density, drink photo density, and menu photo density, and restaurant niche width and their interactions. The control variables are restaurant price, age, longitude and latitude. Given that the location is related with customers' tastes and dish preference, I control them to somewhat rule out these noise. The operationalization, descriptive statistics and correlations of variables are displayed in Tables 8,9,10.

Table 8. Measurement of Variables

Variable Type	Variable Name	Measures
Dependent Variable	Restaurant reputation	Star rating(1-5) of a restaurant
Independent Variables	Photo density	Total number of all the photos / total number of reviews of a restaurant
	Inside photo density	Total number of inside photos / total number of reviews about a restaurant
	Outside photo density	Total number of outside photos / total number of reviews of a restaurant
	Food photo density	Total number of food photos / total number of reviews of a restaurant
	Drink photo density	Total number of drink photos / total number of reviews of a restaurant
	Menu photo density	Total number of menu photos / total number of reviews of a restaurant
	Restaurant niche width	Number of cuisines a restaurant occupies
Control Variables	Restaurant age (weeks)	Number of weeks lapsed since a restaurant's first review posted on Yelp.com
	Restaurant longitude	Geographical longitude of a restaurant
	Restaurant latitude	Geographical latitude of a restaurant
	Restaurant price	Price level ranging from $, $$, $$$ to $$$$

Table 9. Descriptive Statistics of Variables

Variables	Mean	Std. Dev	Min	Max
Restaurant reputation	3.587	0.671	1.000	5.000
Photo density	0.122[4]	0.150	0.004	6.667
Inside photo density	0.022	0.054	0.000	1.000
Outside photo density	0.025	0.062	0.000	1.333
Food photo density	0.014	0.038	0.000	1.000
Drink photo density	0.011	0.035	0.000	1.333
Menu photo density	0.020	0.077	0.000	5.333
Restaurant niche width	3.114	1.216	1.000	6.000
Restaurant age (weeks)	262.923	140.620	1.143	584.977
Restaurant longitude	-96.396	28.344	-115.352	8.549
Restaurant latitude	37.703	5.687	32.877	55.991
Restaurant price	1.679	0.645	1.000	4.000

[4] In addition to reviews with one or more photos, there still exist a large number of purely textual reviews without any photos attached in my research dataset. And in this study, photo density is measured as the percentage of the number of photos to the number of reviews of a restaurant. Then it is not surprising the mean of photo density is less than 1. In the future, I would like to give specific focus on the reviews with photos and see whether the findings are still robust.

Table10. Correlations of Variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Reputation	1											
2. Photo	0.033***	1										
3. Inside	0.012	0.396***	1									
4. Outside	-0.050***	0.444***	0.034***	1								
5. Food	0.058***	0.356***	0.026**	-0.007	1							
6. Drink	0.040***	0.276***	0.008	0.013	0.011	1						
7. Menu	0.044***	0.568***	0.016*	-0.012	0.145***	-0.003	1					
8. Price	0.053***	-0.017*	-0.027**	-0.044***	0.052***	0.035***	-0.027***	1				
9. Niche width	0.041***	0.005	-0.016*	-0.010	-0.041***	0.081***	-0.029***	0.057***	1			
10. Age	-0.127***	-0.312***	-0.169***	-0.109***	-0.138***	-0.136***	-0.120***	0.085***	-0.012	1		
11. Longitude	0.117***	0.132***	0.088***	0.122***	0.027***	0.027***	0.003	0.167***	0.021*	-0.051***	1	
12. Latitude	0.125***	0.176***	0.107***	0.118***	0.066***	0.062***	0.017*	0.198***	0.003	-0.083***	0.882***	1

***p<0.001; **p<0.01; *p<0.05

3.4.2 Estimation and Results

3.4.2.1 Model Specification

$$Y_i = \alpha_0 + \alpha_1 X1_i + \alpha_2 X2_i + \alpha_3 Interactions_i + \alpha_4 X3 + \varepsilon_i$$

Y_i is the star rating of restaurant i;

$X1_i$ is a matrix of variables about photos of restaurant i, including overall photo density, inside photo density, outside photo density, food photo density, drink photo density and menu photo density per review;

$X2_i$ is a the niche width of restaurant i;

$Interactions_i$ is a matrix of two-way interaction terms, including $\mathrm{Photo_i} * \mathrm{Nichewidth_i}$, $\mathrm{Inside_i} * \mathrm{Nichewidth_i}$, $\mathrm{Outside_i} * \mathrm{Nichewidth_i}$, $\mathrm{Food_i} * \mathrm{Nichewdith_i}$, $\mathrm{Drink_i} * \mathrm{Nichewidth_i}$, $\mathrm{Menu_i} * \mathrm{Nichewidth_i}$.

$X3_i$ is a matrix of control variables on restaurant i, including age, price, longitude, latitude of restaurant i.

3.4.2.1 Estimation Results

My empirical model is estimated in SAS using Proc Genmod. To clarify my interaction effects, I use stepwise regression with three blocks of variables: controls (Model 1), linear effects (Model 2), and interaction effects (Model 3). As shown in Table 11, I first estimate Model1. The control effects are consistent with the prior literature. Restaurant age significantly negatively affects restaurant reputation (β=-0.001, p<.001) (Sridhar & Srinivasan 2012) while restaurant price significantly increases reputation (β=0.044, p<.001).

Model 2 validates the linear effects of photo density, inside photo density, outside photo density, food photo density, drink photo density and menu photo density, and niche width on restaurant reputation. The coefficients of photo sharing (β=-0.180, p<.05) is significant and negative, thereby suggesting that the growth of restaurant e-WOM via photo sharing is predicted to be harmful. Furthermore, the coefficient of outside photo sharing (β = -0.652, p<.001) is significant and negative, thereby indicating that the more exposure of a restaurant's outside situations (e.g., patio, entrance, outdoor) hurts the restaurant reputation. In contrast, the coefficients of food photo density (β = 0.723, p<.001), drink photo density (β = 0.459, p<.05) and menu photo density (β = 0.435, p<.001) are significant and positive, thereby indicating that the more exposure of a restaurant's food, drink and menu improves the restaurant reputation. Besides, the coefficient of niche width (β = 0.020, p<.001) is significant and positive, thereby indicating that the more cuisines a restaurant masters, the higher its reputation will be.

Model 3 tests the interaction effects. The coefficient of the interaction between photo density and niche width (β = 0.120, $p<.01$) is significant and positive. On the contrary, the coefficients of the interaction between niche width and food photo density (β = -0.369, $p<.05$), and drink photo density (β = -0.363, $p<.01$) are significant and negative.

Table 11. Estimation Results

Variables	Model 1	Model 2	Model 3
Model 1: Control Effects			
Restaurant age	-0.001^{***}	-0.001^{***}	-0.001^{***}
Restaurant longitude	0.001^{**}	0.001^{***}	0.002^{***}
Restaurant latitude	0.008^{***}	0.008^{***}	0.006^{**}
Restaurant price	0.044^{***}	0.035^{***}	
Model 2: Linear Effects			
Photo density		-0.180^{*}	-1.426^{***}
Inside density		-0.082	0.283
Outside density		-0.652^{***}	-0.384
Food density		0.723^{***}	2.132^{***}
Drink density		0.459*	2.757^{***}
Menu density		0.435^{***}	1.571^{***}
Restaurant niche width		0.020^{***}	0.097^{***}
Model 3: Interaction Effects			
Photo*Nichewidth			0.120^{**}
Inside * Nichewidth			0.035
Outside * Nichewidth			-0.023
Food * Nichewidth			-0.369^{*}
Drink * Nichewidth			-0.363^{**}
Menu * Nichewidth			-0.039

***$p<0.001$; **$p<0.01$; *$p<0.05$

3.5 Discussion

3.5.1 Summary of Results

Academic research on the integration of photo & text in e-WOM context is scarce. This study seeks to understand which kind of layout, sequence and density of photos relative to text is more effective for facilitating customers' elaboration and pleasantness feeling of e-WOM, as well as the corresponding attitude towards the product/service. The empirical results provide significant evidence for my research model and hypotheses.

In the experimental study, my results show that (1) in general, text-first presentation is more beneficial in enhancing perceived diagnosticity and the corresponding customers' evaluations the product/service than the one with photo first; (2) separate layout helps more in customers' diagnosticity perceptions and downstream evaluations than alternate layout, especially, these effects are more obvious for photo-first presentation than the text-first presentation; (3) in contrast, alternate layout is better than separate layout in inducing pleasantness in the e-WOM understanding, regardless of the sequence of text and photo. These results suggest that the addition of photo may not always create instrumental effect, depending on the information presentation (e.g., layout, sequence). The results of this experimental study provide several theoretical and managerial implications.

In the field study, I examine the interaction effects of the photo density in a review and organizational niche width on restaurant reputation. I find that (1) the density of photos in a review systematically hurts restaurant reputation, which may be owing to too many photos shared on the outside environment of a restaurant; in contrast, sharing photos on food, drink and menu of a restaurant increases its reputation. I conjecture that the extremely dominating density of photos especially outside photos in a review attenuated the balance between cognitive elaboration elicited by texts and affective imagination incited by photos, as a result, decreasing the interpretation of the restaurant and its important offerings such as food and drink. However, I find that (2) for the restaurant as a generalist occupying multiple cuisines, the more photos shared in a review, the better a restaurant reputation will be. I conjecture that when a restaurant spans multiple cuisines, it becomes difficult for customers to evaluate and then they demand additional easy-processing cues. Thus, photo sharing as a vivid form of e-WOM is able to expose the performance of a restaurant clearly and directly, as a result, simplifying the assessment of a restaurant satisfying various preferences. In contrast, I also find that (3) for the restaurant as a specialist occupying few cuisines, the more photos shared on food and drink in a review, the better a restaurant reputation will be. I conjecture that for a specialized restaurant focusing on single cuisine, sharing photos on food and drink exemplifies the representative and irreplaceable food & drink of the certain cuisine and then amplifies the performance of the restaurant. These results provide several theoretical and practical implications.

3.5.2 Theoretical Implications

First, my findings extend cue-summation theory by providing some insight into the efficiency and ease of cue summation from a channel-switching perspective. Current cue-summation theory mainly holds two arguments: one highlights the quantity of cues involved in the cue-summation, asserting that the more channels involved, the better communication will be achieved (Jiang & Benbasat 2007a); while the other one turns to the relevancy of cues, explaining that given the superiority of multichannel communication to single channel communication, cues in multichannel have to be relevant in order to work up to its advantage (Dimoka et al. 2012). However, these two arguments are basically developed from a static scope that ignores the dynamic process of channel-switching, which is definitely the central component of cue summation theory. My findings related to relative layout and sequence of photo to text vividly depicts when channel-switching between picture and text is efficient—i.e., separate layout with photo first or generally text first displayed and when channel-switching is pleasant—i.e., alternate layout regardless of the channel appearing sequence.

In addition, my work contributes to the literature on e-WOM by exploiting a new form of e-WOM—i.e., user-generated photo. Existing studies on e-WOM mainly examined numeric rating and textual content (Mudambi & Schuff 2010; Yin et al. 2014). My findings extend the literature by incorporating the impacts of PHOTO on consumer information processing and decision making. This research also adds to the literature on multimedia communication (Kim & Lennon 2008; Dimoka et al. 2012; Jiang & Benbasat 2007a) by exploiting the presentational integration of photo & text. This study is the first to examine the impacts of the relativity of the layout and sequence of photo to text on consumer information processing.

3.5.3 Practical Implications

Moreover, this research also provides guidelines for business to leverage pictures in social media to achieve strategic goals. With respect to experimental study, first, my findings highlight the importance of photos in social media on consumer information processing and decision making. Given that the presentation (e.g., layout, sequence) of the pictures in e-WOM may result in positive or negative outcomes for businesses, particularly new businesses that lack well-defined reputations can derive implications for the presentation and sorting strategies of customer opinions from this research. Second, my research suggests that text displayed first acts as a guide improving the elaboration of the whole review and generating favorable attitude while photo first functions as a trouble-maker adding noise and confusion to the elaboration and inducing unfavorable attitude towards product. Given the critical importance of information elaboration for third-party review platforms such as Yelp.com and Openrice.com, managers may be able to add "sequence of text & photo" in their sorting or filtering systems to let the most easy-to-read reviews reach potential customers, as a result, increasing the sales or ROI of ventures on the platforms. Third, I find that separate layout drags up review diagnosticity compared to the alternate layout, especially when photo is first presented. In contrast, alternate layout brings more pleasantness feelings to understand the review than separate layout. The opposite effects of layout in diagnosticity and pleasantness feeling which are two key intermediate outcomes of final decision making implies distinct strategies for advertising hedonic and utilitarian product/service. Concretely speaking, to advertise hedonic products (e.g., game videos), retailers may choose alternate layout with intertwined text and picture to induce more product-inconsistent pleasantness feelings, while for utilitarian products, they better adopt separate layout with text and picture separate framed in order to achieve more accurate understanding of the product.

For the field study, it points out the inconsistent impacts of photo density on restaurant reputation, that is, photo sharing on the outside of a restaurant mitigates its overall reputation while food, drink and menu photos shared in review improve the reputation; Thus, e-WOM platforms such as Yelp.com and Openrice.com can derive implications for the presentation and sorting strategies of e-WOM from this research. For example, Yelp.com can consider the photo factor in updating its default sorting mechanism by putting reviews with more food, drink and menu photos on the top of a business's information page except for the fixed features (e.g., date, rating) to enhance the restaurant reputation. Besides, my findings suggest that the impact of photo density varies by restaurant's niche width. Thus, e-WOM platforms can encourage photo sharing for a restaurant spanning multiple cuisines regardless of photo content while put more weight on the sharing of particular food and drink photos for a specialized restaurant. Finally, my findings provide implications for managers of restaurants in selection of cues (e.g., food or decorations) to expend marketing budgets for the maximum return of investment (ROI).

3.5.4 Limitations and Future Research

Photo e-WOM is far from fully explored in this research. In terms of experimental study, it can be strengthened in several ways in future research. First, different from traditional mediums using pictures taken by professionals, most e-WOM pictures are taken and posted by users, and thus of varied visual quality. Given that high-quality pictures can vividly represent product while low-quality pictures may present objects in an ambiguous way, the visual quality of pictures has to be taken into account for the elaboration on the combined text & photo. Therefore, the impact of additional photos in varied quality is empirically important to answer. Second, distinct from traditional advertising releasing relatively few and carefully timed images, pictures in social media appear in abundance at with great frequency. Moreover, authors of online reviews enjoy the freedom of writing their comments of various lengths—from simple words/phrases to long paragraphs and posting any number of photos they like in their reviews. Thus, future research may provide some insight into the impacts of word length of text and number of photos.

Concerning the field study, it can be strengthened in several ways in the future research. First, I only examined restaurant reviews in Las Vegas. The generalizability of my findings to other contexts demands further empirical studies. Future research needs to consider other business categories (e.g., hotel, beauty) and extends to different cities. Second, I only focus on one organizational characteristic—i.e., niche width, other characteristics such as ownership (family-owned vs. corporate-owned), scale (regional vs. global) call for future investigation. Third, regarding the characteristics of photo e-WOM, I only examine the density of photos in a review but didn't explore the influences of the visual quality of photos such as colorfulness, brightness and contrast which play a crucial role in the representation of organizational performance via photos. Regarding the robustness checks of the findings on photo density, I would like to do additional analysis in the future to rule out possible alternative explanations.

Conclusion

This dissertation sought to understand the role of e-WOM on consumer information processing and behavior in online consumer websites. Three studies were developed to address the impacts of text e-WOM, vote e-WOM and photo e-WOM on consumer psychology and behavior.

Particularly, in the first study, I investigated the joint effects of review, reviewer, and organizational characteristics on review usefulness by building on dual-process theory. Obtained by utilizing restaurant reviews from Yelp.com with ZINB regression, my empirical results suggest that the certainty-embedded review receives fewer usefulness votes when written by a popular reviewer followed by many fans than when written by a less popular reviewer. Furthermore, restaurant niche width magnifies the usefulness of the certainty-embedded review by a popular reviewer while mitigating the usefulness of the certainty-embedded review by an expert reviewer, thereby supporting my hypotheses. My findings help extend the literature on online customer reviews in multiple aspects.

In the second study, I theoretically articulated how online users respond to prior users' helpfulness votes in e-commerce context. My results suggest that the processing of helpfulness votes follows a sequence, i.e., ratio will be utilized first and magnitude will be used later when users have the motivation and ability to internalize it as reference in decision making. I suppose that helpfulness vote ratio can directly function as trust cues, whereas helpfulness magnitude partially acts as trust cues regardless of review valence and type. In terms of manipulating social influence, I employ an accurate factor, i.e., ratio of helpfulness votes to illustrate the valence of social influence, rather than adopting positive or negative opinions of others to reflect the valence of social influence or simply viewing up-votes as a positive influence and down-votes as a negative influence. Regarding the consequences of social influence, I move one step from primitive voting behavior of subsequent users to a more commercial purchasing behavior.

In the third study, I examined which kind of layout, sequence and density of photos relative to text is more effective for facilitating customers' elaboration and pleasantness feeling of e-WOM, as well as the corresponding attitude and behavioral evaluation towards the product/service. For photo presentation, my research suggests that text displayed first acts as a guide improving the elaboration of the whole review and generating favorable attitude while photo first functions as a trouble-maker adding noise and confusion to the elaboration and inducing unfavorable attitude towards product. These findings related to relative layout and sequence of photo to text vividly depicts when channel-switching between picture and text is efficient—i.e., separate layout with photo first or generally text first displayed and when channel-switching is pleasant—i.e., alternate layout regardless of the channel appearing sequence. With respect to the photo density, photo sharing on the outside of a restaurant mitigates its overall reputation while food, drink and menu photos shared in review improve the reputation. My work contributes to the literature on e-WOM by exploiting a new form of e-WOM—i.e., user-generated photo.

Researchers have long advocated the integration of IS and marketing studies (Bharadwaj et al. 2007; Talvinen 1995; Ozimec et al. 2010). While e-WOM has been discussed in the previous literature, different forms of e-WOM, ranging from textual message, to numeric votes and finally to visual photos, remain largely unexplored. This dissertation is a pioneering effort and contributes to the knowledge of e-WOM in both theoretical and practical aspects. In the future, I will illustrate the potency of field experiment and big data to help tease out the subtle but inevitable effects of specific e-WOM forms such as voice, video, and photo, which are in the rocket trip with the growth of sharing economy. For example, I intend to answer these important questions—i.e., how expert reviews influence marketers' knowledge accumulation in the short- and long-run, and what is the impact of phonic or video e-WOM in the service providing contexts (e.g., game-playing and virtual product purchase).

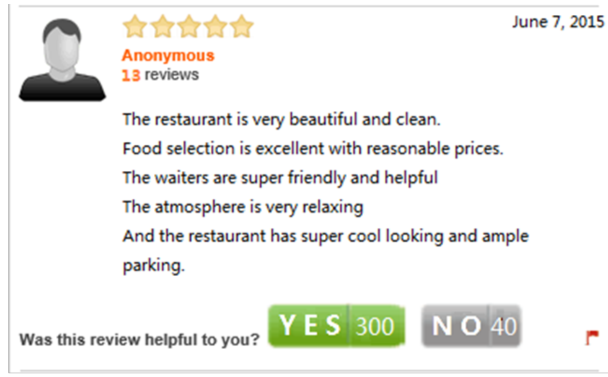
Appendix A—Test Hypotheses Using the Dataset 6 Months before Data Collection (Study 1)

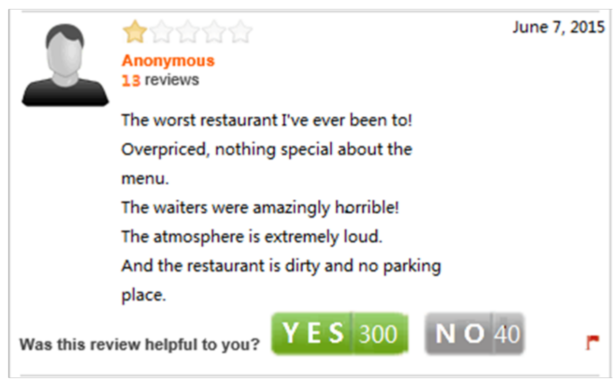
Variables	NB Model Estimating No. of Useful Votes			
	Model 1	Model 2	Model 3	Model 4
Model 1: Control effects				
Rating	-0.4620***	-0.6246***	-0.6272***	-0.6400***
Squared rating	0.0511***	0.0756***	0.0761***	0.0793***
Length	0.0003*	0.0004**	0.0004**	0.0004***
Anger	-0.0015	0.0003	0.0014	00030
Anxiety	-0.0208	-0.0140	-0.0164	-0.0178
Positivity	0.0011	0.0031	0.0029	0.0021
Negativity	0.0111	0.0087	0.0087	0.0085
Readability	0.0079**	0.0060*	0.0060*	0.0052
Log(Review Timespan)#	1	1	1	1
Reviewer Status	1.4382***	1.0099***	1.0116***	0.6605***
Reviewer Yelping Time	0.0003*	0.0002	0.0002	0.0001
Reviewer Average Rating	0.364	0.0576**	0.0585**	0.0613**
Restaurant Reputation	0.2711***	0.3038***	0.3012***	0.2852***
Restaurant Popularity	0.0001	0.0002	0.0002	0.0002
Restaurant age	-0.0014***	-0.0013***	-0.0013***	-0.0013***
Price: $	-0.2124	-0.0609	-0.0785	-0.0613
Price: $$	-0.0682	0.0884	0.0675	0.0879
Price: $$$	-0.0461	0.0606	0.0397	0.0780
Price: $$$$	0.0000	0.0000	0.0000	0.0000
Model 2: Linear effects				
Review Certainty		0.0109	0.0118	0.0126
Reviewer Expertise		0.0002*	0.0002	0.0007***
Reviewer Popularity		0.0183***	0.0179***	0.0416***
Restaurant Niche Width		-0.0416**	-0.0413**	-0.0380*
Model 3: Two-way Interaction effects				
Certainty*Expertise (**H1**)			0.0001	0.0001
Certainty*Popularity (**H2**)			-0.0027*	-0.0014
Certainty*Niche width				0.0178
Expertise*Popularity				-0.0000***
Expertise* Niche width				-0.0000
Popularity* Niche width				-0.0012
Model 4: Three-way Interaction effects				
Certainty*Expertise*Niche width (**H3**)				-0.0002*
Certainty*Popularity*Niche width (**H4**)				0.0050***

Appendix B—Experiment 1 Scenarios (Study 2)

Appendix C—Experiment 2 Scenarios (Study 2)

Appendix D—Experiment 3 Scenarios (Study 2)





Appendix E1— Sample Characteristics in Experimental Study (Study 3)

Variable	**Category**	**Freq.**	**Percent**
Gender	Male	95	58.6
	Female	67	41.4
Education	Did not finish high school	1	0.6
	High school graduate or some degree	52	32.1
	College graduate	75	46.3
	Postgraduate degree	34	21
Income	<$15,000	21	13
	$15,001-$25,000	30	18.5
	$25,001-$35,000	31	19.1
	$35,001-$50,000	30	18.5
	$50,001-$75,000	29	17.9
	$75,001-$100,000	16	9.9
	$100,001-$150,000	5	3.1
	>$150,000	0	0
Japanese food liking	Like a great deal	58	35.6
	Like a moderate amount	60	36.8
	Like a little	34	20.9
	Neither like nor dislike	8	4.9
	Dislike a little	2	1.2
	Dislike a moderate amount	0	0
	Dislike a great deal	0	0
Japanese restaurant experience	Yes	149	92
	No	13	8

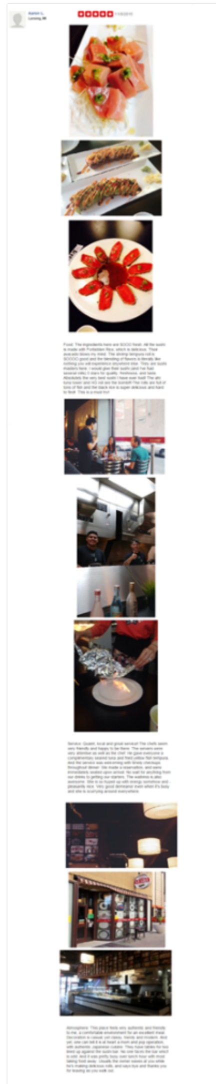
Appendix E2—Measurement in Experimental Study (Study 3)

Constructs	Measurement	Source
Attitude	What do you think of this product/brand? 1. Bad/good 2. Unfavorable/favorable 3. Dislike/like	Rucker & Petty (2004)
Diagnosticity	1. The review helped me familiarize myself with the product. 2. The review helped me evaluate the product. 3. The review helped me understand the performance of the product.	Kempf & Smith (1998)
Pleasantness	1. The review makes me feel happy/unhappy (R). 2. The review makes me feel annoyed/pleased. 3. The webpage makes me feel satisfied/unsatisfied (R). 4. The review makes me feel melancholic/contented. 5. The review makes me feel hopeful/despairing(R). 6. The review makes me feel uncomfortable/comfortable	Mehrabian &Russell (1974)

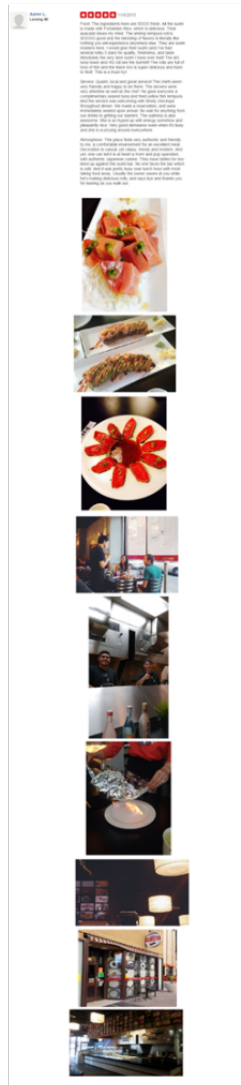
Appendix E3—Experimental Scenarios in Experimental Study (Study 3)

Alternate-UGT first

Alternate-UGP first



Separate-UGT first



Separate-UGP first

REFERENCES

Abelson RP .1988. Conviction. *American Psychologist* (43:4), pp. 267-275.

Ackerloff, G. 1970. “The Market for Lemons: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics* (84:3), pp. 488-500.

Allison PD .2012. “Logistic Regression Using SAS: Theory and Application,” SAS Institute.

Anderson, L. R., and Holt, C. A. 1997. “Information Cascades in the Laboratory,” *American Economic Review* (87:5), pp. 847-862.

Angst, C. M., & Agarwal, R. 2009. “Adoption of Electronic Health Records in the Presence of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion,” *MIS Quarterly* (33:2), pp. 339-370.

Arndt, J. 1967. “Role of Product-Related Conversations in the Diffusion of a New Product,” *Journal of Marketing Research* (4:3), pp. 291-295.

Aydin, T. O., Smolic, A., and Gross, M. 2015. “Automated Aesthetic Analysis of Photographic Images,” Visualization and Computer Graphics, *IEEE Transactions on Visualization and Computer Graphics* (21:1), pp. 31-42.

Ba S., and Pavlou P. A. 2002. “Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior,” *MIS Quarterly* (26:3), pp. 243-268.

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), pp. 99-126.

Bharadwaj, S., Bharadwaj, A., & Bendoly, E. 2007. “The Performance Effects of Complementarities between Information Systems, Marketing, Manufacturing, and Supply Chain Processes.” *Information Systems Research* (18:4), pp. 437-453.

Biswas D, Biswas A, and Das, N .2006. “The Differential Effects of Celebrity and Expert Endorsements on Consumer Risk Perceptions. The Role of Consumer Knowledge, Perceived Congruency, and Product Technology Orientation,” *Journal of Advertising* (35:2), pp. 17-31.

Burnkrant RE and Cousineau A (1975) “Informational and Normative Social Influence in Buyer Behavior.” *Journal of Consumer Research* (2:3), pp. 206-215.

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), pp.

511-521.

Chen, P.-Y., Dhanasobhon, S., and Smith, M. D. 2008. “All Reviews Are Not Created Equal: The Disaggregate Impact of Reviews and Reviewers at Amazon. Com,” Working Paper, Carnegie Mellon University (available at: SSRN: http:// ssrn.com/ abstract =918083).

Cheung, C. M., & Thadani, D. R. (2012). “The Impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model.” *Decision Support Systems* (54:1), pp. 461-470.

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.

Chu SC and Kim Y. 2011. “Determinants of Consumer Engagement in Electronic Word-of-Mouth (eWOM) in Social Networking Sites.” *International Journal of Advertising* (30:1), pp. 47-75.

Clarkson JJ, Tormala ZL, and Rucker DD .2008. “A New Look at the Consequences of Attitude Certainty: the Amplification Hypothesis.” *Journal of Personality and Social Psychology* (95:4), pp. 810-825.

Coxe S, West SG, and Aiken LS (2009) “The analysis of count data: a gentle introduction to Poisson regression and its alternatives,” *Journal of Personality Assessment* (91:2), pp.121-136.

Cyr, D., Head, M., Larios, H., and Pan, B. 2009. “Exploring Human Images in Website Design: A Multi-Method Approach,” *MIS Quarterly* (33:3), pp. 539-566.

Dellarocas, C., Zhang, X. M., 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), pp. 23-45.

Deng, L., and Poole, M. S. 2010. “Affect in Web Interfaces: A Study of the Impacts of Web Page Visual Complexity and Order,” *MIS Quarterly* (34:4), pp. 711-730.

Deutsch, M., and Gerard, H. B. 1955. “A Study of Normative and Informational Social Influences upon Individual Judgment,” *Journal of Abnormal and Social Psychology* (51:3), p. 629.

Dimoka, A., Hong, Y., and Pavlou, P. A. 2012. “On Product Uncertainty in Online Markets: Theory and Evidence,” *MIS Quarterly* (36:2), pp. 395-426.

Duan, W., Gu, B., and Whinston, A. 2009. “Informational Cascades and Software Adoption on the Internet: An Empirical Investigation,” *MIS Quarterly* (33:1), pp. 23-48.

Edell, J. A., and Staelin, R. 1983. "The Information Processing of Pictures in Print Advertisements," *Journal of Consumer Research* (10:1), pp. 45-61.

Fang, Y. H. 2012. "Does Online Interactivity Matter? Exploring the Role of Interactivity Strategies in Consumer Decision Making," *Computers in Human Behavior* (28:5), pp. 1790-1804.

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.

Gardner W, Mulvey EP, and Shaw EC. 1995. "Regression Analyses of Counts and Rates: Poisson, over-dispersed Poisson, and Negative Binomial Models." *Psychological Bulletin* (118:3), pp. 392-404.

Ghose A, and Ipeirotis PG. 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.

Godes, D., and Mayzlin, D. 2004. "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science* (23:4), pp. 545-560.

Goh, K.-Y., Heng, C.-S., and Lin, Z. 2013. "Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User-and Marketer-Generated Content," *Information Systems Research* (24:1), pp. 88-107.

Graziano, W. G., Jensen-Campbell, L. A., Shebilske, L. J., and Lundgren, S. R. 1993. "Social Influence, Sex Differences, and Judgments of Beauty: Putting the Interpersonal Back in Interpersonal Attraction," *Journal of Personality and Social Psychology* (65:3), pp. 522.

Grifoni, P. 2009. Multimodal Human Computer Interaction and Pervasive Services. IGI Global.

Hayes, A. F. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach. Guilford Press.

Holbrook, M. B. 1978. "Beyond Attitude Structure: Toward the Informational Determinants of Attitude," *Journal of Marketing Research* (15:4), pp. 545-556.

Hong W, Chan FK, Thong JY, Chasalow LC, and Dhillon G. 2013. "A Framework and Guidelines for Context-specific Theorizing in Information Systems Research." *Information Systems Research* (25:1), pp.111-136.

Hsu G, Hannan MT, Koçak O .2009. “Multiple Category Memberships in Markets.” *American Sociology Review* (74:1), pp.150–169.

Humphreys A .2010. “Megamarketing: The Creation of Markets as a Social Process.” *Journal of Marketing* (74:2), pp.1-19.

Iturrioz J, Azpeitia I, and Díaz O .2014. “Generalizing the Like Button: Empowering Websites with Monitoring Capabilities.” In Proceedings of the Twenty-Ninth Annual ACM Symposium on Applied Computing, ACM Publications, Gyeongju, Korea.

Jiang, Z., and Benbasat, I. 2007a. “The Effects of Presentation Formats and Task Complexity on Online Consumers' Product Understanding,” *MIS Quarterly* (31:3), pp. 475-500.

Jiang, Z., and Benbasat, I. 2007b. “Research Note-Investigating the Influence of the Functional Mechanisms of Online Product Presentations,” *Information Systems Research* (18:4), pp. 454-470.

Jones Q, Ravid G, and Rafaeli S. 2004. “Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration.” *Information Systems Research* (15:2), pp.194-210.

Kahneman, D., and Tversky, A. 1979. “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica: Journal of the Econometric Society* (47:2), pp. 263-291.

Karmarkar UR, and Tormala ZL .2010. “Believe Me, I Have No Idea What I’m Talking about: The Effects of Source Certainty on Consumer Involvement and Persuasion,” *Journal of Consumer Research* (36:6), pp.1033-1049.

Kelman HC .1961. “Processes of Opinion Change.” *Public Opinion Quarterly* (25:1), pp.57-78.

Kempf, D. S., and Smith, R. E. 1998. “Consumer Processing of Product Trial and the Influence of Prior Advertising: A Structural Modeling Approach,” *Journal of Marketing Research* (35:3), pp. 325-338.

Kim, M., and Lennon, S. 2008. “The Effects of Visual and Verbal Information on Attitudes and Purchase Intentions in Internet Shopping,” *Psychology & Marketing* (25:2), pp. 146-178.

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.

Kotler, P., Keller, K. L., Ancarani, F., & Costabile, M. (2014). Marketing Management 14/e. Pearson.

Kovács, B., Carroll, G. R., and Lehman, D. W. 2013. "Authenticity and Consumer Value Ratings: Empirical Tests from the Restaurant Domain," *Organization Science* (25:2), pp. 458-478.

Kuan, K. K., Zhong, Y., and Chau, P. Y. 2014. "Informational and Normative Social Influence in Group-Buying: Evidence from Self-Reported and EEG Data," *Journal of Management Information Systems* (30:4), pp. 151-178.

Lee, J., Park, D.-H., and Han, I. 2008. "The Effect of Negative Online Consumer Reviews on Product Attitude: An Information Processing View," *Electronic Commerce Research and Applications* (7:3), pp. 341-352.

Lee, K., Lee, B., & Oh, W. 2015. "Thumbs Up, Sales Up? The Contingent Effect of Facebook Likes on Sales Performance in Social Commerce," *Journal of Management Information Systems* (32:4), pp.109-143.

Liu Y, Huang X, An A, and Yu X .2008. "Modeling and Predicting the Helpfulness of Online Reviews." In Proceedings of Eighth IEEE International Conference on Data Mining, IEEE Computer Society, Pisa, Italy.

Liu, Y. 2006. "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing* (70:3), pp. 74-89.

Luca M .2011. "Reviews, Reputation, and Revenue: The Case of Yelp.com," HBS Working Paper 12-016, Harvard Business School, Boston.

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., and Pfann, G. 2013. "More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates," *Journal of Marketing* (77:1), pp. 87-103.

Maddux JE, and Rogers RW .1980. "Effects of Source Expertness, Physical Attractiveness, and Supporting Arguments on Persuasion: A Case of Brains over Beauty," *Journal of Personality and Social Psychology* (39:2), pp. 235-244.

Martin, W. C., and Lueg, J. E. 2013. "Modeling Word-of-Mouth Usage,"*Journal of Business Research* (66:7), pp. 801-808.

Mayer, R. C., Davis, J. H., and Schoorman, F. D. 1995. "An Integrative Model of Organizational Trust," *Academy of Management Review* (20:3), pp. 709-734.

Mehrabian, A., and Russell, J. A. 1974. *An Approach to Environmental Psychology*. the MIT Press.

Moe, W. W., and Trusov, M. 2011. "The Value of Social Dynamics in Online Product Ratings Forums,"*Journal of Marketing Research* (48:3), pp. 444-456.

Muchnik, L., Aral, S., and Taylor, S. J. 2013. "Social Influence Bias: A Randomized Experiment," *Science* (341:6146), pp. 647-651.

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

Ngo-Ye TL and Sinha AP .2014. "The Influence of Reviewer Engagement Characteristics on Online Review Helpfulness: A Text Regression Model," *Decision Support Systems* (61), pp. 47-58.

Otterbacher J .2009. "Helpfulness in Online Communities: a Measure of Message Quality," In Proceedings of the Twenty-seventh SIGCHI Conference on Human Factors in Computing Systems, ACM Publications, Boston, USA.

Ozimec, A. M., Natter, M., & Reutterer, T. 2010. "Geographical Information Systems–based Marketing Decisions: Effects of Alternative Visualizations on Decision Quality," *Journal of Marketing* (74:6), pp. 94-110.

Pan Y and Zhang JQ .2011. "Born Unequal: a Study of the Helpfulness of User-generated Product Reviews," *Journal of Retailing* (87:4), pp. 598-612.

Park, D.-H., and Lee, J. 2009. "Ewom Overload and Its Effect on Consumer Behavioral Intention Depending on Consumer Involvement," *Electronic Commerce Research and Applications* (7:4), pp. 386-398.

Pennebaker JW, Booth RJ, and Francis ME (2007) Linguistic inquiry and word count (LIWC2007), Austin, TX: LIWC (http://www.liwc.net).

Peterson, M., and McGee, S. 1974. "Effects of Imagery Instructions, Imagery Ratings, and Number of Dictionary Meanings Upon Recognition and Recall," *Journal of Experimental Psychology* (102:6), pp. 1007-1014.

Petty RE and Cacioppo JT (1986) "Communication and Persuasion: Central and Peripheral Routes to Attitude Change." Springer Science & Business Media.

Purnawirawan, N., De Pelsmacker, P., and Dens, N. 2012. "Balance and Sequence in Online Reviews: How Perceived Usefulness Affects Attitudes and Intentions," *Journal of Interactive Marketing* (26:4), pp. 244-255.

Racherla P and Friske W .2012. "Perceived 'Usefulness' of Online Consumer Reviews: An

Exploratory Investigation Across Three Services Categories," *Electronic Commerce Research and Applications* (11:6), pp. 548-559.

Rao, A. R., and Monroe, K. B. 1989. "The Effect of Price, Brand Name, and Store Name on Buyers' Perceptions of Product Quality: An Integrative Review," *Journal of Marketing Research* (26:3), pp. 351-357.

Resnick P, Kuwabara K, Zeckhauser R, and Friedman E .2000. "Reputation Systems," *Communications of the ACM* (43:12), pp. 45-48.

Rucker, D. D., and Petty, R. E. 2004. "When Resistance Is Futile: Consequences of Failed Counterarguing for Attitude Certainty," *Journal of Personality and Social Psychology* (86:2), pp. 219-235.

Sen S and Lerman D (2007). "Why Are You Telling Me This? An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing* (21:4), pp. 76-94.

Senecal, S., and Nantel, J. 2004. "The Influence of Online Product Recommendations on Consumers' Online Choices," *Journal of Retailing* (80:2), pp. 159-169.

Severin, W. 1967. "Another Look at Cue Summation," *AV Communication Review* (15:3), pp. 233-245.

Sniezek JA and Van Swol LM .2001. "Trust, Confidence, and Expertise in a Judge-Advisor System," *Organizational Behavior and Human Decision Processes* (84:2), pp. 288-307.

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

Talvinen, J. M. 1995. "Information Systems in Marketing: Identifying Opportunities for New Applications," *European Journal of Marketing* (29:1), pp. 8-26.

Tirunillai, S., & Tellis, G. J. (2012). "Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance," *Marketing Science* (31:2), pp.198-215.

Wei L, Xu W, and Islands C .2014. "Exploring Heuristic Cues for Consumer Perceptions of Online Reviews Helpfulness: The Case of Yelp.com." In Proceedings of the Nineteenth Pacific Asia Conference on Information Systems (Keng Siau, Qing Li, and Xunhua Guo Eds), Chengdu, China.

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), pp. 539-560.

Zeithaml VA .1988. “Consumer Perceptions of Price, Quality, and Value,” *Journal of Marketing* (52:3), pp. 2–22.

Zhang, J., and Liu, P. 2012. “Rational Herding in Microloan Markets,” *Management Science* (58:5), pp. 892-912.