Understanding the roles of forum search aids in consumer information search and response to online product reviews

Guan, Chong

2013

Guan, C. (2013). Understanding the roles of forum search aids in consumer information search and response to online product reviews. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/54872>

<https://doi.org/10.32657/10356/54872>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**UNDERSTANDING THE ROLES OF FORUM SEARCH AIDS
IN CONSUMER INFORMATION SEARCH AND RESPONSE
TO ONLINE PRODUCT REVIEWS**

**GUAN CHONG
NANYANG BUSINESS SCHOOL
2013**

**UNDERSTANDING THE ROLES OF FORUM SEARCH AIDS
IN CONSUMER INFORMATION SEARCH AND RESPONSE
TO ONLINE PRODUCT REVIEWS**

GUAN C.

GUAN CHONG

Nanyang Business School

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2013

ACKNOWLEDGEMENTS

Many people contributed to this dissertation in innumerable ways, and I would like to acknowledge the support and guidance of all of them. First and foremost, I will like to thank my supervisor Prof. Lam Shun Yin. He helped me out a lot in terms of providing much needed critical input to my research and supported me in every way I needed throughout the candidature. I am exceedingly grateful of his generosity with his time and advice. Without his support, this project would not have been possible. I would also like to thank Prof. Naresh K. Malhotra, Prof. Chung Tuck Siong and Prof. Zhang Xiaoling for their invaluable feedback for my dissertation as my committee members. Their encouragement and enthusiasm were important for the completion of this project. Prof Wee Chow Hou, Head of Division, played different roles in providing enlightenment, and allowing me to continue on my path. I would like to thank him for his faith in me, without which, I would not have come so far. Also, I am very appreciative of the other faculty members in the Marketing and International Business division, who provided valuable discussions and training during my graduate studies. The Nanyang Technological University, the Institute on Asian Consumer Insight (ACI) and the SIM University generously provided financial support for this research and I am grateful of their support. I should not forget to thank the Ph.D office and the division secretary Hasnah. They were always very helpful and understanding. Of course, I could not finish this candidature without the support from my family and

friends. Finally, I will like to thank the examiners for sharing their valuable insights and constructive comments on how I can improve on my dissertation.

ABSTRACT

In online product review forums, consumers often encounter a vast number of reviews under a threaded discussion. As processing all the reviews is both nearly impossible and unlikely to be worth the effort and time, it gives rise to important concerns about how consumers select particular reviews to process and how they process them. Their information search activity, which indicates such selection and processing of reviews, would in turn influence their response to the review information in terms of attitude towards the product under review, evaluation confidence and purchase intention regarding the product, etc. To address this issue, I introduce and test a behavioral model of consumer online review search that incorporates consumers' anchoring bias in information search stage and memory-based decisions during product evaluation stage. I propose that these two mechanisms can manifest themselves in the types of reviews inspected as well as depth of processing, and in the decision making process respectively. Forum search aids (summary statistics) are proposed to influence consumer information search behaviors in the selection and processing of different reviews and thereby the response to the product under review. Specifically, online reviews are classified as reviews that are consistent with an average rating (consistent reviews) and those that deviate from the average rating (deviant reviews). The search aids are expected to exert different effects on the search variables corresponding to these two types of reviews.

The results from an in-depth interview and an eye-tracking experiment support many of the proposed relationships, and, in particular, confirm the mediating role of search variables. In the search stage, the results show that presence of an average rating has a positive effect on the two major components of search time, namely, the preview ratio and the average search depth of consistent reviews. Presence of an average rating also has a positive effect on the average search depth of deviant reviews, owing to the distinctiveness of these reviews. Additionally, the effect of distribution presence on the preview ratio of consistent reviews is negative, on condition that the average rating is provided. In the evaluation stage, search time spent on both consistent and deviant reviews significantly affect their product attitude. Across the two stages, provision of an average rating affects product attitude through search variables only when the product rating distribution is absent and when the average rating is low. On the other hand, the provision of a product rating distribution shifts the attitude in a direction opposite to the valence of the average rating through the search variables. The findings provide insights into the complex relationships between search aids, search variables, product attitude and purchase intention during both search and evaluation stages. Finally, managerial implications of this work are described, and potential extensions of the joint model are suggested.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION	10
CHAPTER 2 CONSUMER SEARCH IN ONLINE REVIEW FORUMS	17
Prior Work on Online Review Forums	17
Online Review Search Motivation.....	28
Theories Relevant to Consumer Search in Online Review Forums	29
Economics of Information	30
Anchoring Effect.....	32
Accessibility - diagnosticity Theory	33
The Expanded Framework for Analyzing Review Search	37
Information Search Stage.....	40
Impact of the Provision of an Average Rating on Preview Ratio.....	40
Impact of the Provision of a Rating Distribution on Preview Ratio	42
Levels of the Average Rating	44
Average Search Depth	46
Product Evaluation Stage.....	49
Search Time	49
CHAPTER 3 EMPIRICAL MODEL.....	55
Submodel A	59
Submodel B	65
CHAPTER 4 STUDY 1	69
Methods.....	69
Operationalization of the Review Classification	73
Summary of Results of Study 1	78
Data Coding	78
Summary of Qualitative Findings	81
General Online Review Search Behaviors.....	81
Search Process on the Two Websites.....	84
Eye Movement Data Analysis	90
Conclusion	91

CHAPTER 5 STUDY 2	93
Methods.....	93
Experimental Design.....	93
Measurements of Search Behaviors.....	96
Materials	98
Procedure	100
Qualitative Analysis of Interviews in Study 2	103
Purpose for Conducting Interviews in Study 2	103
Data Collection	103
Developing Coding Categories	104
Linking Specific Quotes to Identified Categories.....	106
Summary of Results	106
Search Aids	107
Evaluation Strategies	116
Feelings	118
Conclusion	119
Results of the Eye Movement Data Analysis	120
Data Preparation and Review Classification.....	120
Measurement Model	121
Structural Equation Model and Research Hypotheses.....	130
Submodel A	131
Submodel B.....	133
Joint Model	137
Effects of Average Rating Presence on Preview Ratios	145
The Effects of Average Rating Presence on Average Search Depth	148
Effects of the Distribution Presence on Preview Ratios and Average Search Depth.....	148
Effects of the Preview Ratios and Average Search Depth on Search Time .	150
Search Time $\times$ Level of the Average Rating Interaction Effect on Product Attitude	151
Product Attitude on Purchase Intention	152

Model Validation	154
CHAPTER 6 DISCUSSION.....	159
Managerial Implications	166
Limitations and Future Research	171
Appendix A Interview Guide for Study 1	176
Appendix B	178
Study 2 Guide	178
Samples Websites Based on the Manipulations.....	187
Appendix C Survey Items	189
Appendix D	192
Appendix E	194
REFERENCE.....	195

CHAPTER 1

INTRODUCTION

Online peer-to-peer review forums have recently become a popular medium that facilitates sharing of huge amounts of user-generated contents. Research demonstrates that online reviews increase consumer interest in the product under discussion (Bickart and Schindler 2001) and greatly speed up the rate of adoption of an innovation (Rogers 2003). The impact of online product reviews on consumers and marketers have attracted the attention of both researchers and practitioners (Bickart and Schindler 2001; Trusov, Bucklin, and Pauwels 2009; Xia and Bechwati 2008).

Major research investigating the impact of online reviews has examined three main aspects: the reviewers, the review messages transmitted, and the information receivers' influence on sales or consumer purchase decisions. When only very few reviews (e.g., one or two reviews) are shown, certain characteristics of the review messages, reviewers and the recipients are found to affect the purchase decision (Dellarocas, Zhang, and Awad 2007; Zhang and Watts 2008; Zhu and Zhang 2010). Interestingly, the impact of voluminous reviews varies across studies. Some found product reviews (e.g., volume, overall product ratings) to influence product sales, while others did not (Chevalier and Mayzlin 2006; Godes and Mayzlin 2004; Liu 2006; Trusov et al. 2009). This implies that a potentially more complicated search

and decision making process is driving consumer response to large amount of word-of-mouth (WOM) opinions under one threaded discussion.

Owing to the large amount of review information under each discussion, processing every review is both nearly impossible and unlikely to be worth the effort and time. To reduce consumers' cognitive effort, many review forums provide online information search aids to aggregate all the review information under a discussion. Although considerable research has examined the impact of online reviews, little effort has been made to examine the influences of search aids on the consumer information gathering processes leading to consumer response in the current research context. According to behavioral decision research, people select information acquisition strategies that are contingent with the decision contexts, in order to achieve a certain level of decision accuracy while saving substantial effort (Payne and Bettman 1992). Extending this area of research, it is proposed that in online review forums, depending on the search aids provided, consumers may allocate their cognitive resources (e.g., search time) strategically on certain reviews. The summary statistics may affect the selection of such information gathering strategies and influence the search behaviors by affecting the perceived benefit of certain reviews. That is, some consumers may rely on summary statistics to select the relevant reviews to process instead of scanning each review on their own. This is important because the information gathered can

play an essential role in shaping consumers' attitude towards the focal product under review.

In sum, information search process in the online review context constitutes a new topic that is worthy of investigation. In the search literature, previous work on economics of information focusing on normative models of optimal search behavior has received much support (Putrevu and Ratchford 1997; Ratchford and Srinivasan 1993; Stigler 1961); whereas behavioral decision research has also produced empirical evidence of departures from the decision rules prescribed by such normative models (Häubl, Dellaert, and Donkers 2010; Zwick et al. 2003). Building on the background of both normative and nonnormative determinants of search behavior in different bodies of research, this research aims at developing and verifying an integrative research model of consumer online review search that extends a baseline sequential information search based on the economics of information theory by incorporating nonnormative influences. The proposed nonnormative influences consist of consumers' anchoring bias on the average rating in information search stage and memory-based decisions (depending on the accessibility of the information) during the product evaluation stage.

The main objective of this research is to explain and predict how the presence of certain search aids (e.g., an average rating and/or a rating distribution) influences the search of reviews in the search stage by extending the anchoring literature, such as the anchoring effect of average rating presence and the forewarning effect of

distribution presence. The two major streams of research on anchoring and accessibility-diagnostics adapted to investigate the impact of the two search aids form the central contribution and arguments of the research.

Extending the anchoring (Epley and Gilovich 2006) and the accessibility-diagnostics (Feldman and Lynch 1988) literature, I propose that these two nonnormative mechanisms can manifest themselves in the type of the reviews inspected as well as depth of processing, and in how the information gathered influences the decision making process respectively. Specifically, I posit that provision of average rating anchors consumers in selecting reviews with ratings close to the average rating (consistent reviews) and reduces their selection of reviews that are inconsistent with the average rating (deviant reviews). In contrast, I argue that provision of frequency distribution forewarns consumers of the dispersion of consumers' rating, thus diverting their selection to the deviant reviews and away from the consistent reviews. Additionally, the average rating could also serve as a salient theme that consumers can organise the processed consistent reviews around and thus increase the impact of consistent reviews on product evaluation.

A series of studies, including a qualitative study (Study 1) and an eye-tracking experiment (Study 2) are conducted to formulate and test the model. In the eye-tracking experiment, eye movement data are gathered to capture the search time allocation. In the current research, I look into product-related consumer

responses as the outcome variables, as they are of great interest to the product marketers. From the forum owner's perspective, it is also acknowledged that consumer review search behaviors can influence their responses to the review forum (e.g., search experience in the forum and revisiting intention) as well. However, considering the limit of time and scope of the research, I focus on the product-related response as the outcome of consumer online review search behaviors. More discussion on forum-related outcome variables is included in the future research section.

This paper seeks to make a theoretical contribution by positing how consumers proceed with processing large amounts of online reviews and explaining why they process and respond to the information in a particular pattern using an integrated behavioral model that extends a baseline normative search process with specific nonnormative influences. The search variables are empirically captured and analyzed based on consumers' eye movements over a threaded discussion. Most prior WOM or online review research were based on self-reported attitude/intention data or actual sales data, which did not capture the underlying process regarding how consumers search these reviews. The current study measures the consumers' eye movements while processing the reviews, including the exact area that the consumer is paying attention to and the amount of attention paid while processing the reviews. This process tracing method is important as it enables accurate measurements of the number of reviews previewed, the number of

reviews selected to read and the time spent reading them within each set of reviews. Thus, the findings can add new knowledge to online review research. A better understanding of consumer search strategies is also managerially important because it is during information acquisition that marketers influence consumers' buying decisions (Schmidt and Spreng 1996). Thus, the practical implications include pragmatic guidelines for marketing managers to provide responses to certain reviews selectively and to provide certain search aids in the interface so as to improve consumers' brand or product perceptions.

Besides contributing to the search in online reviews literature, I extended the anchoring and accessibility-diagnostics literature to explain and predict how the presence of certain search aids (e.g., an average rating and/or a rating distribution) influences the search of reviews in the search stage. To my knowledge, this research is the first to theoretically and empirically analyze the role of search aids on online consumer review search and thereby product evaluation.

The remainder of the paper is organized as follows. In the next section, I briefly review research that investigates the impact of online reviews. This review highlights some inconsistent findings, suggesting the need to consider the underlying search process for voluminous reviews. Then, I introduce how consumers proceed in information search in online forums in the absence of search aids as a baseline search process building on the economics of information literature. Subsequently, I review the theories that form the conceptual model,

namely the anchoring (Epley and Gilovich 2006) and the accessibility-diagnosticsity (Feldman and Lynch 1988) literature. I then develop hypotheses about how forum search aids exert effects on search behaviors and consumer response subsequently. This is followed by this research's methodology, results and discussion. Lastly, the managerial implications and future research are presented.

CHAPTER 2

CONSUMER SEARCH IN ONLINE REVIEW FORUMS

Prior Work on Online Review Forums

Online review forums are Internet platforms that allow consumers to read articulations of other consumers and share their own opinions, experiences, and related knowledge etc., on a great number of goods, services, and companies (Hennig-Thurau and Walsh 2003). In the context of consumer reviews, information content is contained in the main texts (some containing photos, e.g., Tripadvisor.com) of the reviews. A review is a single opinion referring to a specific situation and time. Owing to the length of the reviews and the large amount of reviews under a discussion, most review forums only show the title, posting date and a preview of the first few sentences of a message. Consumers may click on “read more” (hyperlinked) to access a full message or “read less” to go back to the preview. Each review is also succinctly summarized in the form of an overall numeric product rating indicating its valence and extremity. A rating appears prominently at the top of the review, and usually ranges from one to five stars or ten stars (see Figure 1 for an example of individual reviews and corresponding ratings). In some forums, ratings on certain product attributes are also provided. There are usually ten reviews (in preview format) appearing on each page (e.g., Tripadvisor.com and hungrygowhere.com) and consumers can click on the page numbers to access reviews on other pages.

Figure 1

Examples of Individual Reviews and Corresponding Ratings

(a) a review message on Tripadvisor.com

http://www.tripadvisor.com/Hotel_Review-g294264-d1447339-Reviews-Hard_Rock_Hotel_Singapore-Sentosa_Island.html - Retrieved on 2011-03-06.

Ti Chan
Singapore,
Singapore
81 contributions

“Nice Modern/Spacious Rooms, Bad Views. not a good base to explore SG”

Reviewer ratings for this hotel:

Value	Service
Rooms	Sleep Quality
Location	
Cleanliness	

Date of review: Feb 26, 2011

Stayed here during office event so I didn't pay myself but it's about 209++ USD/night (without breakfast). If you want to focus on visiting places in Sentosa (and get away from the city), this is a good place. But if you want to visit Singapore as a whole - this is NOT a good base (Sentosa you can easily commute...
[more](#)

(b) a review message on hungrygowhere.com

http://www.hungrygowhere.com/singapore/miss_u_cafe/ - Retrieved on 2011-03-06.

eastieboy Recommend.

Food/Beverage: 9 Ambience: 9 Value: 8 Service: 9

8.8 Overall

East Side Pride
Mar 6, 2007

Finally a cosy place to chill out in this eastern region of Singapore. This cafe serves surprisingly tasty delights at affordable prices. The menu holds a wide variety of appetizers, italian cuisine (includes pizzas), western food (includes grilled gourmets). I only managed to try the Fried Calamari which was above average with the fritto sauce, i tried the salami pizza which was delicious, however the Salmon Salad had more salmon than veggies in this salad! What i also noticed about [\(Read More\)](#)

A summary of major research investigating the impact of online reviews is shown in Table 1. The table shows that previous studies on the impact of online reviews have examined three main aspects: the reviewers, the review messages transmitted, and the information receivers' influence on sales or consumer purchase decisions. Furthermore, these empirical studies can be classified into two

categories: individual-level analysis and market-level analysis. Individual-level studies mainly focus on the impact of one or very limited number of reviews and examine the specific characteristics of the reviewers, review messages, and the information receivers' influence on consumer purchase decisions. Table 1 shows that at an individual consumer level, review messages containing arguments that are of high quality and that confirm prior beliefs tend to have a stronger impact on purchase decisions. Reviewer characteristics, such as credibility, identity disclosure, similarity with the review viewer and certain review recipient characteristics such as prior knowledge and product involvement can moderate the effect of the online reviews. However, at the market level, findings on the impact of the reviews are less consistent. The market-level studies use aggregated panel data on the characteristics of the reviews (e.g., the average ratings, volumes and the valence of consumer reviews) extracted from the websites to examine the impact of review messages on product sales. Some found product reviews (e.g., volume, overall product ratings) to influence product sales, while others did not (Chevalier and Mayzlin 2006; Godes and Mayzlin 2004; Liu 2006; Trusov et al. 2009). The overall pattern of the findings suggests that when only very few reviews are shown, certain characteristics of the review messages, reviewers and the recipients affect the purchase decision, whereas the impact of voluminous reviews varies across studies. This implies that a potentially more complicated search and decision making process is driving consumer response to large amount of word-of-mouth

(WOM) opinions under one threaded discussion. Because the search process data is rarely collected, the consumer information gathering processes leading to consumer response to the reviews have not been well investigated.

In searching for information about the product under review, consumers are unlikely to read every review and spend the same amount of time on each review that they choose to read. As the selection and time spent on processing the reviews potentially affect their evaluation of the product, it is important to understand what affect these search variables. In this paper, I argue that provision of product rating statistics (search aids) by forum owners affects these variables and thereby consumers' attitude towards the product under review. In the remaining sections of this chapter, I would first introduce the online review search aids and the search motives. Subsequently, I review the anchoring (Epley and Gilovich 2006) and the accessibility- diagnosticity (Feldman and Lynch 1988) literature that form the bases for the conceptual model. Then, I introduce how consumers search in online forums in the absence of search aids as a baseline search process. Hypotheses about how forum search aids exert effects on search behaviors and consumer response are then developed subsequently.

Table 1
A Summary of Major Research on Factors Influencing Online Reviews Effectiveness

Factors Influencing Online Review Effectiveness	Author and Year	Online Review Effectiveness Measurements	Findings
Reviewer Characteristics			
Source Credibility -Expertise -Trustworthiness	Cheung et al. (2009), Hu, Liu, and Zhang (2008), Zhang and Watts (2008)	Product attitude and purchase intention, online product choice, attitude toward the Web site and credibility of the eWOM messages (Cheung et al. 2009) Sales (Hu et al. 2008) Information adoption (Zhang and Watts 2008)	Source credibility affects perceived credibility of the review, which affects the attitude, purchase intention, and product choice (Cheung et al. 2009). The reputation and exposure of sources positively affect product sales (Hu et al. 2008). Messages perceived to have higher source credibility produce a higher level of information adoption (Zhang and Watts 2008).
Disclosure of identity	Forman, Ghose, and Wiesenfeld (2008), Hu et al. (2008)	Sales (Forman et al. 2008)	The disclosure of reviewer identities positively affects sales (Forman et al. 2008; Hu et al. 2008).
Social tie	Steffes and Burgee (2009)	Intention to choose (Steffes and Burgee 2009)	Strong tie referrals are more influential than weak tie ones in affecting intention to choose (Steffes and Burgee 2009).
Similarity -in age, gender, education, social status, and geographical location	Steffes and Burgee (2009) Forman et al. (2008)		Similarities in education and social statuses positively affect the intention to choose (Steffes and Burgee 2009). Reviewers from the same geographical location as the readers affect their credibility positively, which in turn affect sales positively (Forman et al. 2008).
Gender and age of reviewers	Dellarocas, Zhang, and Awad (2007)	Sales (Dellarocas et al. 2007)	The gender and age of online reviewers affect sales (Dellarocas et al. 2007).

Factors Influencing Online Review Effectiveness	Author and Year	Online Review Effectiveness Measurements	Findings
Review Message Characteristics			
Argument quality - Relevance -Timeliness -Accuracy -Comprehensiveness/ completeness	Cheung et al.(2009), Lee, Park, and Han (2008), Park, Lee, and Han (2007), Sher and Lee (2009), Zhang and Watts (2008)	Product attitude (Lee et al. 2008; Sher and Lee 2009) Purchase Intention (Park et al. 2007)	Argument strength positively affects review credibility (Cheung et al. 2009), purchase intention (Park et al. 2007), and information adoption (Zhang and Watts 2008). Consumers conform to negative product reviews if the argument quality is high (Lee et al. 2008). Highly skeptical consumers are indifferent to the message quality (Sher and Lee 2009).
Message sidedness	Cheung et al. (2009)	Trust in the eWOM and product attitude (Cheung et al. 2009)	Message sidedness has a weak impact on the product attitude and review credibility (Cheung et al. 2009).
Information that confirms /disconfirms prior beliefs	Cheung et al.(2009), Zhang and Watts (2008)		Reviews that confirm prior beliefs have a positive effect on review credibility and product attitude (Cheung et al. 2009). Argument quality affects information adoption more and source credibility affects it less when the information is highly disconfirming (Zhang and Watts 2008).
Message valence -Absolute valence -Relative valence	Ba and Pavlou(2002), Chatterjee (2001), Cheung et al. (2009), Chevalier and Mayzlin (2006), Clemons, Gao, and Hitt(2006), Dellarocas et al. (2007), Gauri, Bhatnagar, and Rao (2008), Hu et al. (2008), Xia and Bechwati (2008), Chen, Liu and Zhang (2011)	Price premiums and buyer behaviour (Ba and Pavlou 2002) Retailer evaluation and patronage intention (Chatterjee 2001) Sales (Clemons et al. 2006) Online Store Loyalty (Gauri et al. 2008)	The number of positive reviews induces trust in seller's credibility and in turn gives higher price premiums (Ba and Pavlou 2002). Message valence influences the attitude of review readers (Cheung et al. 2009), relative sales (Chevalier and Mayzlin 2006) and actual sales (Dellarocas et al. 2007; Hu et al. 2008). The most positive quartile of reviews affects sales positively (Clemons et al. 2006). Positive reviews have the greatest impact on repurchase intention (Gauri et al. 2008). The effect of cognitive personalization on purchase intention is moderated by valence (Xia and Bechwati 2008). Reviews exert significant impact on stock returns in the direction of their valence relative to other previously published reviews (Chen, Liu, and Zhang 2011).

Factors Influencing Online Review Effectiveness	Author and Year	Online Review Effectiveness Measurements	Findings
Message length	Chevalier and Mayzlin (2006)		Longer reviews do not necessarily stimulate sales (Chevalier and Mayzlin 2006).
Review contents - experiential vs. factual -attribute-centric vs. benefit-centric	Xia and Bechwati (2008)		The nature of product reviews affects purchase intention (Xia and Bechwati 2008).
Overall rating	Chatterjee (2001), Cheung et al. (2009), Chevalier and Mayzlin (2006), Duan et al. (2008), Godes and Mayzlin (2004), Liu(2006), Zhu and Zhang (2010)	Sales (Chatterjee 2001; Chevalier and Mayzlin 2006; Duan et al. 2008; Godes and Mayzlin 2004; Liu 2006; Zhu and Zhang 2010)	The overall rating affects sales positively (Chatterjee 2001; Chevalier and Mayzlin 2006; Duan et al. 2008; Godes and Mayzlin 2004). Ratings contribute to the product attitude and review credibility (Cheung et al. 2009). An increase in overall rating results in higher incremental sales for products (Zhu and Zhang 2010). Online ratings have no impact on sales (Duan et al. 2008; Liu 2006).
Volume (Total number of reviews)	Chevalier and Mayzlin (2006), Davis and Khazanchi (2008), Dellarocas et al. (2007), Duan et al. (2008), Gauri et al. (2008), Godes and Mayzlin (2004), Lee et al. (2008), Liu (2006), Park et al. (2007), Sher and Lee (2009), Trusov et al. (2009)	Sales (Davis and Khazanchi 2008) Percentage of households that watched the show (Godes and Mayzlin 2004) Revenue (Liu 2006; Trusov et al. 2009)	The volume of reviews does not affect sales (Chevalier and Mayzlin 2006), customer loyalty (Gauri et al. 2008), and percentage of households that watched the show (Godes and Mayzlin 2004). Pure increase in review volume has no effect on sales (Davis and Khazanchi 2008). The volume of reviews has a positive impact on revenue (Duan et al. 2008; Liu 2006; Trusov et al. 2009), and purchasing intention (Park et al. 2007). The early volume of online reviews can be used as a proxy of early sales (Dellarocas et al. 2007). The proportion of negative reviews in the total number of reviews results in a change in product attitude (Lee et al. 2008). Consumers with low skepticism tend to form attitude peripherally and are more persuaded by review quantity (Sher and Lee 2009).

Factors Influencing Online Review Effectiveness	Author and Year	Online Review Effectiveness Measurements	Findings
Consistency (the extent to which a review is congruent with others' opinions)	Cheung et al.(2009), Clemons et al. (2006)		Consistency of reviews has a positive effect on review adoption (Cheung et al. 2009) and sales (Clemons et al. 2006).
Dispersion (the extent that the conversations take place across many communities)	Dellarocas et al., (2007), Godes and Mayzlin (2004)		The dispersion of reviews has a positive influence on future sales (Dellarocas et al. 2007). Dispersion influences the percentage of households that watch the show more than volume (Godes and Mayzlin 2004).
Review Information Receiver Characteristics			
Prior knowledge (of the topic /the platform)	Cheung et al. (2009), Chatterjee (2001), Zhu and Zhang (2010)		Confirmation of prior belief has a positive effect on perceived review adoption (Cheung et al. 2009). Prior knowledge of the product increases the effect of overall rating of product on sales (Zhu and Zhang 2010).
Involvement	Cheung et al. (2009), Gruen et al.(2006), Lee et al. (2008), Park et al. (2007)		Involvement in the reviews has a positive effect on product attitudes (Cheung et al. 2009). High-involvement consumers tend to conform in attitude changes more than low-involvement consumers (Lee et al. 2008). Low-involvement consumers are affected by the quantity rather than the quality of reviews, but high-involvement consumers are affected by review quantity mainly when the review quality is high (Park et al. 2007).
Cognitive personalization (processing the information in a self-referential manner)	Xia and Bechwati (2008)		The level of cognitive personalization affects purchase intention (Xia and Bechwati 2008).

Factors Influencing Online Review Effectiveness	Author and Year	Online Review Effectiveness Measurements	Findings
Others			
Nature of the product (experience vs. search)	Huang et al. (2009), Park et al. (2009), Senecal and Nantel (2004), Xia and Bechwati (2008)	Consumer search (time per page and total number of pages) and purchase behaviour (Huang et al. 2009) Online product choices (Senecal and Nantel 2004)	Product reviews have a greater effect on consumer search and purchase behavior for experience than for search goods (Huang et al. 2009; Senecal and Nantel 2004). Product types moderate the effect of eWOM information direction on website reputation (Park et al. 2007). The nature of the product reviewed (experience vs. search) affects purchase intention (Xia and Bechwati 2008).
Product popularity	Zhu and Zhang (2010)		Higher review ratings result in increment of sales if the product is popular already (Zhu and Zhang 2010).
Website's reputation (established vs. unestablished)	Park et al. (2009)		EWOM is more effective for established websites than for unestablished websites (Park et al. 2007).

Online Review Search Aids

Search aids are technologies designed to interact with consumers to help them acquire better information and/or to do so with less effort, similar to decision aids (DAs). Much research on online decision aids has relied on the effort–accuracy framework of cognition to investigate the beneficial impact of decision aids on reducing the cognitive effort expended by users and increasing their decision quality (accuracy) in online shopping context (Diehl, Kornish, and Lynch 2003; Haubl and Trifts 2000). Similar to DAs in online shopping, many review forums also create interfaces with summarising or sorting aids to facilitate consumers' search process, owing to the massive amounts of information under a threaded review discussion. Popular aggregated data includes average product rating, the distribution of product ratings and helpfulness ratings. A distribution of the ratings indicates the frequencies of different ratings' occurrence (see Figure 2 for examples of product rating distributions). In online review forums, when the mean is held constant, the distribution of past ratings reflects the level of consensus among the reviews. Consensus refers to the level of agreement between opinions (West and Broniarczyk 1998). Depending on the level of consensus, the distribution of ratings can vary from a unimodal shape to a bimodal shape. For a unimodal shape distribution, the highest frequencies cluster around the mean of the distribution with a gradual tailing off on either side in the frequency of values as they move away from the center of the distribution. A bimodal distribution has many extreme rating, and very few ratings in between. This paper looks

into situations where the review ratings form a unimodal distribution, as it commonly appears across many product/service categories. Sorting functions usually include sorting by product rating or by helpful votes of the review (see Table 2). The forum search aids share some commonalities with the online shopping decision aids in reducing search effort as a primary motivation. However, the main purpose of search aids used in forums is to assist consumers in information gathering, allowing them faster access to the relevant information, by minimizing the cost in identifying/locating certain reviews or enabling the consumers to estimate the review reservation utilities before reading them.

Table 2

Popular International and Local Review Web Sites and Search Aids

Review Site	Search Aids Provided	
	<i>Summarizing Function</i>	<i>Sorting Function</i>
Amazon.com	Average Rating, distribution Review Helpfulness Voting	Sort by product rating, sort by date, sort by helpfulness of review messages
Eopinions.com	Average Rating Only	Sort by product rating, sort by date
Hungrygowhere.com	Average Rating, distribution Review Helpfulness Voting	Sort by product rating, sort by date, sort by helpfulness of review messages
Sg.movies.yahoo.com	Average Rating Review Helpfulness Voting	Sort by date
Sgcarmart.com	Average Rating	Sort by product rating, sort by date
Tripadvisor.com	Average Rating Review Helpfulness Voting	Sort by product rating, sort by date

Figure 2
Examples of Distributions of Product Ratings

(a) a unimodal distribution	(b) a bimodal distribution
http://www.tripadvisor.com.sg/Hotel_Review-g297697-d308413-Kuta_Paradiso_Hotel_Bali-Kuta_Bali.html	http://www.amazon.com/Britney-Spears-Greatest-Hits-Prerogative/dp/B00064AM62/ref=sr_1_3?s=dvd&ie=UTF8&qid=1301412369&sr=1-3

Online Review Search Motivation

Consumers approach online reviews for a variety of motives, including getting purchase-related information, being part of the community and learning how to use a product, etc (Goldsmith and Horowitz 2006; Hennig-Thurau and Walsh 2003). According to the framework proposed by Moe (2003), shopping motives could be categorized under two dimensions: (a) directed versus exploratory and (b) the potential purchasing horizon: immediate versus future. Following this framework, as online product review forums are essentially a product information search platform, both prepurchase and knowledge building motives may result in future purchase. The difference lies in the first dimension proposed by Moe (2003): whether one's search is goal-directed with a planned purchase in mind or exploratory. The objective of prepurchase visits is to acquire relevant information to help make an optimal choice. Therefore, the

strategy is designed to build the consideration set and evaluate the items in the set. In other words, the searcher is in the market for a particular category of product but is still unsure of which specific product in the category to buy. On the other hand, knowledge building is to acquire a bank of relevant product information potentially useful in the future. In these situations, the search motive is exploratory and the consumer's main objective is to increase product and/or marketplace expertise. The consumer is not necessarily considering any specific purchase, but the information acquired may influence future purchasing decisions. In the current research, I focus on prepurchase scenarios only, as it is the most important and prevailing motive for people to approach reviews online (Goldsmith and Horowitz 2006; Hennig-Thurau and Walsh 2003; Huang and Yang 2008). Thus, the presumption is that the consumer has a planned purchase in mind and is aware of the product under discussion from passive search activities (e.g., exposure to offline WOM), but she has little knowledge about the product and thus unsure about the true product utility.

Theories Relevant to Consumer Search in Online Review Forums

In the current research, I posit that the normative baseline search process largely explains consumer review search behaviors in the absence of search aids based on the economics of information theory (the normative search process). However, introducing forum search aids will bring about specific behavioral influences in both information search stage and product evaluation stage that add to the baseline process. The proposed nonnormative influences consist of consumers' anchoring bias on the average rating and negative bias on a low

average rating (depending on the diagnosticity) in information search stage and memory-based decisions (depending on the accessibility) in the product evaluation stage. Extending the anchoring (Epley and Gilovich 2006) and the accessibility - diagnosticity (Feldman and Lynch 1988) literature, this research posits that these two nonnormative mechanisms can manifest themselves in whichever set of the reviews inspected as well as depth of processing, and in how the information gathered influences the decision making process respectively. I now introduce these relevant streams of research leading to the baseline model - the economics of information, the literature about the effect anchoring and the accessibility-diagnosticity theory.

Economics of Information

The economics of information perspective posits that consumers seek to maximize net benefits from information contingent upon their perception of the search environment, reflected in the trade-off between the perceived benefit and cost of search (Moorthy, Ratchford, and Talukdar 1997; Stigler 1961). Extending the above theory to the current research context, when consumers are faced with online reviews, they adopt search behaviors that are optimal, contingent upon the search aids provided. The optimality of a consumer's search is also reflected in her trade-off between the perceived benefit and cost of search. By reading the reviews, she gets to ascertain the utility of the product under review in order to make a purchase decision. The cost of search includes the time and effort involved in (a) locating certain reviews, depending on the interface provided by the review forum; and (b) reading the reviews.

Whether a consumer will read certain reviews is affected by the reviews' reservation utility - a summary measure of the value of getting information from a particular review, considering the perceived marginal utility and the cost of reading it (Moorthy et al. 1997). A consumer evaluates a product using a utility function. She would first process the review(s) with the highest reservation utility. After each review is read, the mean and/or dispersion of the product utility distribution may change. For instance, reading a persuasive positive review can increase the mean and possibly reduce the dispersion. Subsequently, she continues to search review(s) with the highest reservation utility in the rest of the discussion until the marginal benefit of search equals to the search cost.

Building on the economics of information theoretical backdrop, this paper posits that consumers approach the reviews sequentially (following the default sequence) in the absence of forum search aids. This is based on the three following reasons. First, one does not know the exact distribution or average of the product ratings. Therefore, she is not sure about which reviews are more representative or relevant. She may simply assume reviews are equally relevant across different ratings. Second, even though she may feel certain reviews are more useful based on the valence or extremity indicated by the corresponding ratings (e.g., people may want to search more negative information), identifying and locating these particular reviews from the list requires too much effort. Whereas on the other hand, search cost is lowest for the next review in the sequence. Third, the reviews are listed by date of post with the recent post

appearing first by default. People may consider recent posts more relevant for some product/service categories in which the quality can change over time (e.g., hotels). This default sequence creates an ordered environment based on one fixed strategy: listing the reviews in a descending order of utility (review recency). Assuming the average product ratings and product rating distributions do not vary over time, the reviews that a consumer sequentially processes will follow a similar distribution as the overall discussion.

Starting from the baseline search process presented, I move on to introduce the specific behavioral influences on consumers' search and decision making process that complement the normative effect. The forum search aids influencing the search process and the subsequent responses are discussed in the next section.

Anchoring Effect

Anchoring is a cognitive bias that describes the common human tendency to overly rely, or "anchor," on one piece of arbitrary information when making decisions (Epley and Gilovich 2006). According to the anchoring literature, a provided arbitrary value can activate people to test the hypothesis that the true value is equal to the anchor value (Jacowitz and Kahneman 1995). Testing a hypothesis makes the individuals pay more attention to target features that are like the anchor as compared to features that are different from the anchor (Chapman and Johnson 1999). In addition, studies also find that forewarnings about the anchor can moderate anchoring effects by increasing the amount of effortful adjustment (Epley and Gilovich 2005; Tversky and Kahneman 1974).

Anchoring effect occurs when people place too much importance on one piece of information (the average rating), causing negligence of other probabilities (Chapman and Johnson 1999; Epley and Gilovich 2006).

The current research framework builds upon this stream of research and posits that an average product rating exerts a similar effect as an “anchor” and a product rating distribution could “forewarn” about and thus mitigate such anchoring effect. When an average rating is present and the distribution is absent, the average rating becomes highly salient. Thus, it is proposed that the average rating can exert a similar effect as an “anchor”, activating consumers to test the hypothesis that the product utility equals to this value, and neglect the dispersion of product ratings. The presence of a rating distribution with dispersion is similar to offering a warning about the anchoring effect on the average rating, which can reduce the anchoring bias.

Accessibility - diagnosticity Theory

According to the diagnosticity approach in the impression formation literature, people build an impression of a stimulus using the most diagnostic information available (Skowronski and Carlston 1989). Diagnosticity refers to the sufficiency of the retrieved input to arrive at a solution for the judgment task at hand (Menon, Raghurir, and Schwarz 1995). One of the major factors affecting the diagnosticity of information is its discrepancy from expectancy (Fiske 1980). Furthermore, negative information is generally more diagnostic than positive information and is more useful for forming impressions (Skowronski and Carlston 1989).

Applied to the current context, a consumer that is consistently exposed to products with high average ratings in daily life will view a product with a low average rating as a much more informative/diagnostic “anchor” and the confirmation bias associated with it should increase.

Although search behaviors largely reflect the information consumers gathered, not all the gathered information would be retrieved and used in the decision making process in the same manner. Most judgments that people make about a target object are memory based rather than stimulus based (Feldman and Lynch 1988). That is, they are based on a sampling of thoughts of the object that a perceiver has built up during the entire information search process and stored in memory (Lingle et al. 1979). As proposed by Feldman and Lynch (1988), an earlier piece of information will be used as an input to a subsequent response if the former is accessible and if it is perceived to be more diagnostic than other accessible inputs. In the current research, I focus on the memory-based decision in view of the accessibility of the information in memory. Accessibility refers to the ease of retrieving an input from memory (Menon, Raghubir, and Schwarz 1995). The likelihood that a piece of information will be used in a later judgment task is a positive function of the accessibility of the information in memory (Simmons, Bickart, and Lynch 1993).

Search time allocated to a certain set of reviews will increase the accessibility of the corresponding reviews and therefore the likelihood of this information being used in product evaluation also increases. As a result, their

product attitude will change towards the opinions advocated by these reviews. These effects of search time on attitude could be positive or negative depending on the level of the average rating.

In the next section, I discuss the hypotheses elaborating on the effects of both search aids (the nonnormative search process) are introduced subsequently. Figure 4 presents the overarching theoretical framework for the behavioral model of consumer review search. The following theoretical discussion and hypothesis development are centered on this framework.

Figure 4
Theoretical Framework

Behavioral Model of Consumer Review Search

The Expanded Framework for Analyzing Review Search

Before discussing the specific hypotheses, I now present an overview of the expanded framework for analyzing forum search behaviors. When controlling for the argument quality and length of the review messages, product ratings reflect the valence and the extremity of the corresponding review contents. For the purpose of the current study, I categorize the reviews under a threaded discussion into two sets based on their corresponding product ratings and the average product rating: reviews that are consistent with the average rating (consistent reviews) and reviews that deviate from the average (deviant reviews). To some extent, the average rating represents the group opinion or group consensus, as it is the “averaged” view on the product expressed by the different reviews. Research on group consensus shows that people respond to information that affirms the group opinion and information that disagrees with it differently (Matz and Wood 2005). Extending the findings to the current context, the consumers should also respond differently to reviews that are consistent with the average rating and those that are not. This forms the theoretical basis for the classification. The classification of the reviews is based on the actual average rating of the discussion - the mean of all the review ratings in the threaded discussion. The average rating exists regardless whether the average rating value is shown to the viewers or not. Even when the average rating is not provided to the viewers, they could still form a rough

judgment of where the average rating lies by themselves after reading some of the reviews.

Search time allocation is defined as the total amount of time a consumer devotes into gathering and processing the information from each of the above mentioned sets of reviews in an aggregated sense. This is of great importance because it can affect subsequent decision making. Search time has been identified as a key variable of interest in information search studies (Ratchford and Srinivasan 1993). The current research expands the traditional measures of search time to the current research context and breaks it down to selection and processing.

Within each set of reviews, the selection process consists of two stages: selection of reviews to preview and selection of reviews to read. Selection of reviews to preview consists of two components: reviews available and preview ratio. Preview ratio is the probability of each review being previewed within each set of reviews. This concept corresponds to a consumer's initial decision on whether to allocate attention to a particular set of review information. Selection of reviews to read also consists of two components: number of reviews previewed and the full view ratio. Provided that a review is previewed, full view ratio is the probability of this review being read (clicking on "read more"). Whether a consumer choose to inspect the detailed view of certain reviews can depend on a number of other factors, including informativeness or diagnosticity of the actual review contents (Skowronski and Carlston 1989). In addition, the processing of the reviews depends on the average search depth

within this set of reviews, which refers to the average amount of information gathered from each review read within a particular set. This concept reflects levels-of-processing in information processing theory, falling on a shallow to deep continuum (Craik 2002). A deeper level of search is associated with more attention given, extraction and comprehension of meaning, and elaborated analysis. This decomposition allows a distinction between different search aids in their selection effects, and a more precise estimation of those effects. Figure 3 presents an overview of the expanded framework for analyzing search behaviors within each set of reviews.

Figure 3
An Expanded Framework for Analyzing Search Variables within Each Set of Reviews (Consistent and Deviant)

The anchoring literature is first reviewed and different effects of search aids (summary statistics) in the information search stage are introduced. This is followed by the review of memory-based decision literature based on the accessibility- diagnosticity theory and a discussion on the impacts of the search aids and search time in the product evaluation stage. Hypotheses on how search aids affect search behaviors and assist product evaluation are discussed in turn.

Information Search Stage

Impact of the Provision of an Average Rating on Preview Ratio

In the current context, assuming reviews have already been sorted by date, consumers would search sequentially from top to bottom in the absence of search aids (baseline search process). When an average rating is present and the distribution is absent, the average rating becomes highly salient due to its numerical format and high summarizing power. Thus, the average rating can exert a similar effect as an “anchor”. This will activate consumers to test the hypothesis that the product utility equals to this value, and neglect the dispersion of product ratings. Furthermore, studies have repeatedly found that people prefer to test a hypothesis in a one-sided way, by searching for evidence consistent with the hypothesis they hold at a given time (Kunda 1999; Nickerson 1998). Rather than searching through all the relevant evidence, people tend to gather information that is affirmative to their hypothesis (Baron 2000). This inclination is often called confirmation bias (Jonas, Schulz-Hardt, and Frey 2005; Wason 1968), ascribed to cognitive dissonance theory (Festinger 1957). In particular, when individuals face conflicting cognitions,

they experience an unpleasant feeling, called dissonance. When two cognitions are in direct conflict, anxiety (dissonance) is produced and grows. If the dissonance becomes strong enough, people have a motivational drive to reduce dissonance by adding new ones to create a consistent belief system and reducing the importance of the weaker cognition(s) (Festinger 1957). In this case, individuals prefer to only focus their attention to seek information, and thus generate thoughts, that align with their existing beliefs or attitudes. This type of “positive test” would yield exactly the information supporting the hypothesis. In the current research context, people will more likely inspect a review that is consistent with the average rating and neglect the deviant reviews to confirm their “hypothesis”.

H1(a): Provision of an average rating increases the preview ratio of consistent reviews, when the distribution is absent.

H1(b): Provision of an average rating decreases the preview ratio of deviant reviews, when the distribution is absent.

Anchoring effect occurs when people place too much importance on one piece of information (the average rating), causing negligence of the probability of having a dispersed product rating distribution (Chapman and Johnson 1999; Epley and Gilovich 2006). The presence of a rating distribution with dispersion is similar to offering a warning about the confirmation bias, which can reduce the bias, although not completely eliminate it (Highhouse and Bottrill 1995). Thus, provision of salient information indicating the dispersion of opinions (forewarning the consumers) can mitigate the anchoring effect of the average

rating presence. Applied to the current research, provision of a product rating distribution will decrease the anchoring effect of provision of average rating on the preview ratio of consistent reviews and deviant reviews.

H2(a):Provision of a rating distribution reduces the positive impact of an average rating provision (i.e., making the impact less positive) on the preview ratio of consistent reviews.

H2(b):Provision of a rating distribution reduces the negative impact of an average rating provision (i.e., making the impact less negative) on the preview ratio of deviant reviews.

Impact of the Provision of a Rating Distribution on Preview Ratio

This paper looks into situations where the review ratings form a unimodal distribution with some dispersion (with a few extreme ratings and moderate ratings, both positive and negative), as it commonly appears across many product/service categories. In information processing literature, to account for the phenomena of selective attention, Treisman (1979) suggests that incoming information is subjected to a series of criterion, including salience of the information (e.g., being placed dominantly on a page) and its qualitative nature, such as its contextual relevance (e.g., summarizing power). Salience is defined as distinctiveness, prominence, and obviousness (Treisman 1979). This concept is widely used to refer to any aspect of a stimulus that stands out from the rest. When a product rating distribution with dispersion (with a few extreme ratings and moderate ratings, both positive and negative) is present, the dispersion of opinions becomes especially distinctive. Furthermore, the distribution presence

will increase perceived uncertainty, as it reflects some level of disagreement. When people are confronted with a discrepant message, people will consider this discrepant position in order to understand why some would be willing to deviate (Moskowitz 1996).

When the average rating is absent, for a unimodal distribution with dispersion, the mode of the rating distribution becomes very prominent and relevant. The mode is prominent, as a unimodal distribution has only one peak and the data points are clustered around the mode. In addition, people can form their own judgment of the average rating based on the mode of the distribution, especially when the distribution is unimodal, and the mode of the distribution is consistent with the average rating. Thus, the mode is also considered highly relevant and useful. That is, when the average rating is absent, the mode of the rating distribution is almost as prominent and relevant as the dispersion. Extending the selective attention theory by Treisman (1979), as both the mode and the dispersion of the rating distribution are almost equally prominent and relevant, consumers will not pay more attention to reviews that are consistent with the mode or deviant reviews that are highlighted by the dispersion.

When the average rating is present, the mode of a unimodal rating distribution is less useful as it does not provide any additional information for the viewers. On the other hand, the presence of a numerical average rating with high summarizing power makes the dispersion of the distribution more distinct due to the contrast effect. Consumers would wonder why some are willing to deviate from the average rating. Thus, when the average rating is present, only

the dispersion of the distribution becomes more distinct and relevant. In this case, the presence of a distribution with dispersion highlights the discrepant messages, and directs readers' attention to these deviant reviews and away from the consistent reviews.

To summarize, when the average rating is absent, the mode and the dispersion of the rating distribution are almost equally prominent and relevant. In this case, the distribution presence will not increase consumers' attention to reviews that are consistent with the mode or deviant reviews that are highlighted by the dispersion. Provision of a distribution with dispersion increases the saliency of the deviant reviews, thus consumers are more likely to preview the deviant reviews and less likely to preview the consistent reviews, as compared to the situation where the distribution is absent, only when the average rating is present.

H3: Provision of a product rating distribution does not affect preview ratios of (a) the consistent reviews or (b) the deviant reviews, when the average rating is absent.

H4(a): Provision of a product rating distribution decreases the preview ratio of consistent reviews, when the average rating is present.

H4(b): Provision of a product rating distribution increases the preview ratio of deviant reviews, when the average rating is present.

Levels of the Average Rating

As it was discussed earlier, provision of an average rating as an anchor can lead to a confirmation bias. This confirmation bias can be further

strengthened by a low average rating (e.g., 3.7/10), as compared to a high average rating (e.g., 7.3/10). According to the accessibility - diagnosticity literature, people build an impression of a stimulus using the most diagnostic information available (Skowronski and Carlston 1989). One of the major factors affecting the diagnosticity of information is its discrepancy from expectancy (Fiske 1980). Uncommon information is perceived to be more diagnostic, and therefore has a greater influence on the selection of information, than common information is. Applied to the current context, high average ratings are much more commonly seen on review Web sites, as compared to low average ratings (Hu, Zhang, and Pavlou 2009). Furthermore, negative information is generally more diagnostic than positive information and is more useful for forming impressions (Skowronski and Carlston 1989). A consumer that is consistently exposed to products with high average ratings in daily life will view a product with a low average rating as very unusual and thus very diagnostic. Therefore, she is likely to see a product with a low average rating as even more unsatisfying and is almost sure about her judgment. Additionally, she is unlikely to look through deviant reviews to challenge her current stand. To further ascertain her judgment in order to make a decision, she would still inspect the reviews consistent with the average. That is, a low average rating would be viewed as a much more informative/diagnostic “anchor” and the confirmation bias associated with it should increase.

H5: A low level of average rating magnifies (a) the positive effect of provision of an average rating on the preview ratio of the consistent reviews,

and (b) the negative effect of provision of an average rating on the preview ratio of the deviant reviews.

Average Search Depth

According to the behavioral approach in the information acquisition literature, consumers' information acquisition strategies are adaptive (Payne, Bettman, and Johnson 1988). That is, information is processed in such a way that it approximates the accuracy of normative procedures while saving substantial effort. Extending this theory to the current context, an average product rating can be viewed as a proxy for majority opinion which is highly relevant to the product decision. Presence of such highly issue-relevant information such as an average rating may modify the consumers' information search strategy. That is, consumers adapt their information search strategy based on the average product rating so that they can reach a "good enough" decision while saving much search effort. When a consumer is exposed to some highly issue-relevant information, she may feel that engaging in deep or effortful processing of the review information is not worth the time and effort, as it is unlikely to greatly improve the accuracy of her decision. She may simply scan through the reviews quickly to look for other relevant information that can ascertain her judgment. That is, provision of average product rating can reduce the overall level of processing, which is captured in average search depth in the current context.

H6: Provision of an average rating decreases the average search depth of both (a) consistent and (b) deviant reviews.

Besides the above proposed effects, it is also acknowledged that other alternative effects of the average rating presence on the average search depth might co-exist. According to the anchoring effect, presence of an average rating activates consumers to test the hypothesis that the product utility equals to this value in a one-sided way, by gathering information that is affirmative to the hypothesis (Baron 2000). In this case, in order to confirm the “hypothesis” that the product utility equals to the average rating, individuals also can spend more time on consistent reviews that align with the average rating and spend less time on deviant reviews that may stimulate dissenting beliefs or attitudes.

Furthermore, the presence of an average rating also can exert a positive effect on the average search depth of deviant reviews according to the selective attention phenomena in information processing literature (Treisman 1979). Incoming information is subjected to a series of criterion, such as salience of the information (e.g., being distinctive or obvious) and its qualitative nature, such as its contextual relevance. A deviant review becomes especially distinctive, when the average rating is present, as the average rating can be viewed as a proxy of majority opinion. When one is confronted with a discrepant message, in order to confirm her current judgment based on the majority view, she will examine this discrepant position carefully in order to understand why some would be willing to deviate (Moskowitz 1996). Therefore, the consumer may read the deviant reviews more carefully when she is given the average rating information. Although no formal hypotheses are proposed based on the above two other alternative effects, it is acknowledged

that such alternative effects may cancel out part of the negative impact of average rating presence on the average search depth proposed in H6.

The presence of a product rating distribution with dispersion will increase uncertainty, as it highlights some level of disagreement. Incorrect product attitudes can have deleterious consequences (Petty and Cacioppo 1996), such as a bad holiday experience. To reduce the probability of incorrect product attitudes, consumers are very motivated to reconcile the conflicting views by allocating considerable cognitive effort during the search process. In sum, when a product rating distribution is provided, consumers would carefully examine both consistent and deviant reviews to discover true value of the product.

H7: Provision of a product rating distribution increases the average search depth of both (a) consistent and (b) deviant reviews.

Beside the seven hypotheses proposed above, it is also acknowledged that search aids may also have other secondary impacts on search variables. For example, provision of an average rating may decrease the average search depth of deviant reviews more than that of consistent reviews, as consistent reviews contain more relevant information as compare to the deviant reviews. The reason that these effects are not discussed in this paper is that they are comparatively less important and relevant to the central message that the paper conveys and their effect sizes are expected to be comparatively weaker. Excluding the secondary effects helps to prioritize the main hypotheses, which are more meaningful, plausible and expected to have greater effect size.

Product Evaluation Stage

Higgins (1978) showed that elaborative activities on favourable or unfavourable information selectively about a stimulus caused subsequent evaluations of that stimulus to be either positive or negative, respectively. This is because the elaborative activities performed on the target information during the search stage increases the likelihood that it will be retrieved for use in making a subsequent judgment (Craik 2002; Feldman and Lynch 1988). These elaborative activities include attention and comprehension of the information. Extending the above findings, I argue that search time on a particular set of reviews that captures such elaborative activities can have an impact on the product evaluation. In addition, certain search aid can moderate the evaluation process.

Search Time

When consumers search online reviews, the search time spent on a particular type of reviews (reviews that are consistent with the average rating and reviews that deviate from the average rating) reflects the attention and comprehension of the corresponding reviews. Thus, search time allocated to a certain set of reviews will increase the accessibility of the corresponding reviews and therefore the likelihood this information to be used in product evaluation. As a result, their product attitude will change towards the opinions advocated by these reviews. That is, if a consumer spends more time searching the consistent reviews, her product attitude would become more in line with the average rating. On the other hand, search time on the deviant reviews shifts her

attitude in a direction opposite to the valence of the average rating. These effects of search time on attitude could be positive or negative depending on the level of the average rating.

H8 (a): Search time on consistent reviews increases product attitude when the average rating is high, and decreases product attitude when the average rating is low.

H8 (b): Search time on deviant reviews decreases product attitude when the average rating is high, and increases product attitude when the average rating is low.

Extending the accessibility-diagnosticity model (Feldman and Lynch 1988) to the current research context, certain information gathered becomes more influential when it can be very easily organized around a salient theme (e.g., the average rating) in the memory (accessible). To understand the way in which people make memory-based judgements about a product, it is also necessary to understand how they organize the product-related information in memory and draw on that information in a later decision. Research has demonstrated that when the stimulus information is preceded by a salient theme, the memory for information can be organized around the theme and more easily remembered (Lingle et al. 1979), thus more accessible in subsequent decision making process. Applied to the current study, when an average rating is provided, consumers can organise the processed reviews that are consistent with the average rating around this salient theme. Therefore, provision of an average rating can increase the accessibility of review

information that is congruent with the average rating and the influence of this information on evaluation increases.

H9: Provision of an average rating increases (magnifies) the positive effect of search time on consistent reviews on product attitude when the average rating is high, and the negative effect on product attitude when the average rating is low.

Research on decision aids has found that provision of decision aids improves both objective and subjective decision quality (Haubl and Trifts 2000; Murray and Haubl 2008). This is because decision aids make consumers believe that they are focusing their effort on more promising alternatives in the shopping environment. Similarly, with the help of search aids in review forums, consumers feel that they can easily identify and focus on the more relevant reviews. Furthermore, review forums with well functioning search aids are generally viewed as more professional and trustworthy. In this case, they become more confident about the judgement made on the product under review.

H10: Provision of search aids increases evaluation confidence as compared to the conditions when the search aids are not provided.

Persuasion theory has provided much support for the positive influence of attitude on purchase intentions (Petty and Krosnick 1995). I expect the relationships between attitude and behavioral intention generally hold true in the current research context. As the effect of product attitude on purchase intention is well tested in the literature, no formal hypothesis is set up in this

paper to test this effect. The purpose of excluding this relationship from the main hypotheses is to reinforce the main focus of the paper.

Additionally, research shows that attitude-behavior consistency is moderated by amount of certainty/confidence (Kraus 1995; Petkova, Ajzen, and Driver 1995). Higher confidence is associated with a stronger relationship between attitudes and behavioural intentions (Ajzen 1988). Extending the above finding to the current research context, the product evaluation confidence can strengthen the relationship between attitude difference and purchase intention.

H11: The evaluation confidence strengthens the positive relationship between attitude and purchase intention.

The hypotheses H1 to H11 are summarized in Table 3. Figures 4 and 5 present the theoretical framework and empirical model respectively.

Table 3
Summary of Hypotheses

Information Search Process		Theory Applied
<i>Interaction Effect of Average Rating Presence × Rating Distribution Presence</i>		Anchoring, and Accessibility-Diagnosticity
H1(a)	Provision of an average rating increases the preview ratio of consistent reviews, when the distribution is absent.	
H1(b)	Provision of an average rating decreases the preview ratio of deviant reviews, when the distribution is absent.	
H2	Provision of a product rating distribution decreases (a) the positive impact of an average rating provision (i.e., making the impact less positive) on the preview ratio of consistent reviews, and (b) the negative impact of an average rating provision (i.e., making the impact less negative) on the preview ratio of deviant reviews.	
H3	Provision of a product rating distribution does not affect preview ratios of (a) the consistent reviews or (b) the deviant reviews, when the average rating is absent.	
H4(a)	Provision of a product rating distribution decreases the preview ratio of consistent reviews, when the average rating is present.	
H4(b)	Provision of a product rating distribution increases the preview ratio of deviant reviews, when the average rating is present.	
<i>Interaction Effect of Average Rating Presence × Level of the Average Rating</i>		Accessibility-Diagnosticity
H5	A low level of average rating magnifies (a) the positive effect of provision of an average rating on the preview ratio of the consistent reviews, and (b) the negative effect of provision of an average rating on the preview ratio of the deviant reviews.	
<i>Effect of Average Rating Presence</i>		
H6	Provision of an average rating decreases the average search depth of both (a) consistent and (b) deviant reviews.	
<i>Effect of Rating Distribution Presence</i>		
H7	Provision of a product rating distribution increases the average search depth of both (a) consistent and (b) deviant reviews.	
Product Evaluation Process		
<i>Effects of Search Time Allocation × Level of the Average Rating</i>		
H8(a)	Search time on consistent reviews increases product attitude when the average rating is high, and decreases product attitude when the average rating is low.	
H8(b)	Search time on deviant reviews decreases product attitude when the average rating is high, and increases product attitude when the average rating is low.	
<i>Interaction Effect of Search Time Allocated to Consistent Reviews × Average Rating Presence × Level of the Average Rating</i>		
H9	Provision of an average rating increases (magnifies) the positive effect of search time on consistent reviews on product attitude when the average rating is high, and the negative effect on product attitude when the average rating is low.	
<i>Effect of Search Aids Presence</i>		
H10	Provision of search aids increases evaluation confidence as compared to the conditions when the search aids are not provided.	
<i>Interaction Effect of Attitude × Evaluation Confidence</i>		
H11	The evaluation confidence strengthens the positive relationship between attitude and purchase intention.	

Figure 5
Empirical Framework

CHAPTER 3

EMPIRICAL MODEL

For the purpose of the current study, I categorize the reviews under a threaded discussion into two sets based on their corresponding product ratings and the average product rating: reviews that are consistent with the average rating as the consistent reviews and reviews that deviate from the average as the deviant reviews (see Figure 6). Within each set of reviews, the selection process consists of two stages: selection of reviews to preview and selection of reviews to read. In Figure 6, the number of reviews available is an exogenous variable. Preview ratio is defined as within each set of reviews, provided that a review is present, the probability of this review being previewed. Full view ratio is within each set of reviews, provided that a review is previewed, the probability of this review being read (clicking on “read more”). Average search depth is defined as the level of processing within each set of reviews (average full view time spent per review within each set of the reviews). This decomposition allows a distinction between different search aids in their selection effects, and a more precise estimation of those effects.

Figure 6
An Expanded Framework for Analyzing Search Variables within Each Set of Reviews (Consistent and Deviant)

Following Figure 6, for each set of reviews, the number of reviews previewed is expressed as a product of number of reviews available and the preview ratio, number of reviews read (clicking on “read more”) as a product of the number of reviews previewed and the full view (detailed view) ratio, and total (detailed view) search time allocated as a product of the number of reviews read and average search depth. Thus, total search time spent on a particular type of reviews can be expressed as a product of the number of reviews available, the preview ratio, the full view ratio and average search depth of that type of reviews. That is,

$$T_i = N_{li} \times PR_i \times FVR_i \times ASD_i \quad (1)$$

where

i : an indicator for a particular set of reviews; $i=1, 2$; $i=1$ for consistent review; and $i=2$ for deviant reviews;

N_{li} : number of reviews available for set i ; N_{li} is an exogenous variable;

T_i : total search time allocated to set i ;

PR_i : preview ratio of set i ;

FVR_i : full view ratio of set i ;

ASD_i : average search depth of set i ;

I then take logarithms on both sides of Equation (1), and obtain the following Equation 2:

$$T'_i = N'_{li} + PR'_i + FVR'_i + ASD'_i \quad (2)$$

where $T'_i, N'_{li}, PR'_i, FVR'_i$ and ASD'_i are the log-transform of T_i, N_{li}, PR_i, FVR_i and ASD_i , respectively.

Proposed Model

As shown in Figure 7, the proposed model is decomposed into two submodels (A and B) which consist of different subsets of relationships. Submodel A depicts how the provision of information search aids influences consumer search behaviors during the search stage. Submodel B describes how the search variables affect product evaluation and purchase intentions during the evaluation stage. This

decomposition allows a distinction between the search stage and evaluation stage and simplifies the detailed presentation of the proposed model. In addition, the decomposition also can help isolate potential issues/problems in preliminary data analysis, such as the systematic missing values in the search variables due to incidental truncation, which will be discussed in detail later in the results section. Each of the submodels is introduced in the following sections.

Figure 7
Basic Conceptual Model

Submodel A

Submodel A consists of six hypotheses (H1-H7) as indicated in Figure 8. The submodel and the hypothesized relationships are captured by Equations 2 to 4 as follows.

Figure 8
Expanded Empirical Framework of H1-H7 (Search Stage)

The log-transformed search time spent on a particular type of reviews is expressed as a linear function of the four log-transformed search components in Equation 2 (see p.49 for the derivation of Equation 2).

$$T'_i = N'_{1i} + PR'_i + FVR'_i + ASD'_i \quad (2)$$

where

$i=1, 2$; $i=1$ for consistent review; and $i=2$ for deviant reviews;

$T'_i, N'_{1i}, PR'_i, FVR'_i$ and ASD'_i : the log-transform of T_i, N_{1i}, PR_i, FVR_i and ASD_i , respectively.

The four log-transformed endogenous search variables components, namely, the preview ratios and average search depth of consistent and deviant reviews are expressed as linear functions of their explanatory variables in Equations 3 and 4:

$$PR'_i = \alpha_{1i} + \beta_{1i}ARAT + \beta_{2i}DIST + \beta_{3i}LVAR + \gamma_{1i}ARAT \times DIST + \gamma_{2i}ARAT \times LVAR + \varepsilon_{1i} \quad (3)$$

$$ASD'_i = \alpha_{2i} + \beta_{4i}ARAT + \beta_{5i}DIST + \varepsilon_{2i} \quad (4)$$

where

PR'_i : log-transform of the preview ratio of set i ($i=1, 2$; $i=1$ for consistent review; and $i=2$ for deviant reviews);

ASD'_i : log-transform of the average search depth of set i ;

α_{1i}, α_{2i} : intercepts;

ε_{1i} : disturbance term, assumed to be normally and identically distributed with mean zero and constant variance;

$\beta_{1i}, \beta_{2i}, \beta_{3i}, \beta_{4i}$ and β_{5i} : parameters for effects;

γ_{1i} and γ_{2i} : parameters for interactions;

ARAT, DIST, and LVAR: explanatory variables (ARAT = provision of an average rating (presence=1 and absent =0); DIST = provision of a distribution of product

ratings(presence=1 and absent =0); LVAR= low level of the average rating (low=1 and high=0).

As hypothesized in Table 4, the preview ratio of consistent reviews is explained by the positive effect of average rating presence (ARAT), the negative effect of distribution presence (DIST), the negative interaction effect between average rating presence and distribution presence (ARAT×DIST), and the positive interaction effect between average rating presence and a low average rating level (ARAT×LVAR) (Equation 3; H1a,H2a,H3a, H4a and H5a).

Table 4
Summary of Hypotheses for Submodel A

Information Search Process		Hypotheses Testing
<i>Interaction Effect of Average Rating Presence × Rating Distribution Presence</i>		
H1(a)	Provision of an average rating increases the preview ratio of consistent reviews, when the distribution is absent.	$H1a: \beta_{11} > 0$;
H1(b)	Provision of an average rating decreases the preview ratio of deviant reviews, when the distribution is absent.	$H1b: \beta_{12} < 0$
H2	Provision of a product rating distribution decreases (a) the positive impact of an average rating provision (i.e., making the impact less positive) on the preview ratio of consistent reviews, and (b) the negative impact of an average rating provision (i.e., making the impact less negative) on the preview ratio of deviant reviews.	$H2a: \gamma_{11} < 0$; $H2b: \gamma_{12} > 0$
H3	Provision of a product rating distribution does not affect preview ratios of (a) the consistent reviews or (b) the deviant reviews, when the average rating is absent.	$H3a: \beta_{21} = 0$; $H3b: \beta_{22} = 0$
H4(a)	Provision of a product rating distribution decreases the preview ratio of consistent reviews, when the average rating is present.	$H4a: \gamma_{11} < 0$;
H4(b)	Provision of a product rating distribution increases the preview ratio of deviant reviews, when the average rating is present.	$H4b: \gamma_{12} > 0$
<i>Interaction Effect of Average Rating Presence × Level of the Average Rating</i>		
H5	A low level of average rating magnifies (a) the positive effect of provision of an average rating on the preview ratio of the consistent reviews, and (b) the negative effect of provision of an average rating on the preview ratio of the deviant reviews.	$H5a: \gamma_{21} > 0$; $H5b: \gamma_{22} < 0$
<i>Effect of Average Rating Presence</i>		
H6	Provision of an average rating decreases the average search depth of both (a) consistent and (b) deviant reviews.	$H6a: \beta_{41} < 0$; $H6b: \beta_{42} < 0$
<i>Effect of Rating Distribution Presence</i>		
H7	Provision of a product rating distribution increases the average search depth of both (a) consistent and (b) deviant reviews.	$H7a: \beta_{51} > 0$; $H7b: \beta_{52} > 0$

The preview ratios of consistent and deviant reviews are explained by the same set of factors but in the opposite direction. The preview ratio of deviant reviews is explained by the negative effect of average rating presence (ARAT), the positive effect of distribution presence (DIST), the positive interaction effect between average rating presence and distribution presence (ARAT×DIST), and the negative interaction effect between average rating presence and a low average rating level (ARAT×LVAR) (Equation 3; H1b,H2b,H3b, H4b and H5b).

Full view ratio is the probability of a review being read in detail -clicking on “read more”, provided that this review is previewed. Whether a consumer choose to inspect the detailed view of certain reviews can depend on a number of other factors, including informativeness or diagnosticity of the actual preview contents (Skowronski and Carlston 1989). This is because when a review is previewed by a consumer, it shows that this review is deemed relevant based on the search aids given. To further confirm the selection and to avoid wastage of search effort, she would need to inspect the preview content to ascertain the relevance of the review. Thus, search aids are less influential at this stage.

The average search depth of both consistent and deviant reviews are explained by the negative effect of average rating presence (ARAT) and the positive effect of distribution presence (DIST) (Equation 4; H6 and H7). Control variables include need for closure, need for cognition, knowledge about and involvement in resorts, prior online review experience, and prior attitude. To

emphasize on the hypothesized effects and to simplify the presentations of the equations, these control variables are not presented in the equations. However, they are included in Submodel A in data analysis.

Submodel B

Submodel B consists of four hypotheses (H8-H11) as indicated in Figure 9.

Figure 9
Expanded Empirical Framework of H8-H11 (Evaluation Stage)

For the purpose of this research, logarithmic transformation is used there to represent the curvilinear form of relationships in the respective hypotheses (Lilien, Kotler, and Moorthy 1995). Also, since the search variables are represented as log-transformed variables in earlier equations, the logarithmic transformation has to be preserved for consistency so that all equations in the model can be estimated

simultaneously using SEM. Thus, it is suitable to be used to depict the relationship between search time and attitude in the current context. Prior attitude is included as a covariate in the Submodel B. The relationships between attitude, search time and the level of average rating are expressed as follows.

$$\begin{aligned}
 Att = & \alpha_3 + \beta_1' T_1' + \beta_2' T_2' + \beta_3' LVAR + \beta_4' ARAT \\
 & + \gamma_1' T_1' \times LVAR + \gamma_2' T_2' \times LVAR \\
 & + \gamma_3' T_1' \times ARAT \\
 & + \gamma_4' T_1' \times LVAR \times ARAT + \beta_5' Patt + \varepsilon_3
 \end{aligned} \tag{6}$$

where

T_i : total search time allocated to set i ($i=1, 2$; $i=1$ for consistent review; and $i=2$ for deviant reviews);

T_i' : log-transform of T_i ;

ARAT: provision of an average rating;

LVAR: level of the average rating;

Att: post search attitude;

Patt: prior attitude.

H8 hypothesizes that the presence of an average rating increases the effect of search time of consistent reviews on attitude. As it is hypothesized in H7a, the directions of the impact of search time on attitude depend on the level of the average rating. Thus, based on H7a, H8 hypothesizes a three-way interaction, where the impact of search time could be enhanced by the presence of an average rating and the directions of the impact still depend on the level of the average

rating. Thus γ_4' instead of γ_3' is the parameter estimate testing the relationship hypothesized in H8. Attitude is proposed to positively influence purchase intentions. Additionally, as proposed in H9 and H10, the provision of search aids increases evaluation confidence, which in turn strengthens the influence of attitude on intention (Ajzen 1988). Attitude, evaluation confidence and intention are expressed as linear functions of their explanatory variables, forming the following Submodel B:

$$Con = \alpha_4 + \beta_6' ARAT + \beta_7' DIST + \varepsilon_4 \quad (7)$$

$$Int = \alpha_5 + \beta_8' Att + \beta_9' Con + \gamma_5' Att \times Con + \varepsilon_5 \quad (8)$$

where

ARAT: provision of an average rating;

LVAR: level of the average rating;

Att: post search attitude;

Patt: prior attitude.

Con: evaluation confidence;

Int: purchase intention.

Equations 6-8 constitute the Submodel B. Table 5 summarizes the hypotheses in Submodel B.

Table 5
Summary of Hypotheses for Submodel B

Product Evaluation Process		Hypotheses Testing
<i>Effect of Search Time Allocation × Level of the Average Rating</i>		
H8(a)	Search time on consistent reviews increases product attitude when the average rating is high, and decreases product attitude when the average rating is low.	H8a: $\gamma_1' < 0$
H8(b)	Search time on deviant reviews decreases product attitude when the average rating is high, and increases product attitude when the average rating is low.	H8b: $\gamma_2' > 0$
<i>Interaction Effect of Search Time Allocated to Consistent Reviews × Average Rating Presence × Level of the Average Rating</i>		
H9	Provision of an average rating increases (magnifies) the positive effect of search time on consistent reviews on product attitude when the average rating is high, and the negative effect on product attitude when the average rating is low.	H9: $\gamma_4' > 0$
<i>Effects of Search Aids Presence</i>		
H10	Provision of search aids increases evaluation confidence as compared to the conditions when the search aids are not provided.	H10: $\beta_6' > 0$ and $\beta_7' > 0$
<i>Interaction Effect of Attitude × Evaluation Confidence</i>		
H11	The evaluation confidence strengthens the positive relationship between attitude and purchase intention.	H11: $\gamma_5' > 0$

CHAPTER 4

STUDY 1

A preliminary study (Study 1) was first conducted to confirm the main motivator of online review searching behavior and the type(s) of products/services consumers commonly search reviews for and to refine the conceptual framework and hypotheses for the empirical study. Study 1 is largely qualitative and exploratory in nature. The main data of Study 1 was collected through both concurrent and retrospective think-aloud methods. Then, a major study (Study 2) was conducted to test the hypotheses. Study 2 is mainly quantitative and confirmatory in nature. The main data of Study 2 was collected through both eye-tracking measures and online questionnaires.

Methods

Qualitative methods are valuable in expanding current understanding of consumer behavior (Patton 1990). Think-aloud protocols (TAPs) yield rich, descriptive process data that allows for a deeper understanding of the phenomenon at hand. Very few online review researchers have focused specifically on the search behaviors when consumers approach the reviews. The results of this qualitative study (Study 1) would aid in the refinement of the conceptual framework and hypotheses for the empirical study, and the experimental stimuli and measures designed for Study 2. Thus, Study 1 adopts a think-aloud protocol with eye-tracking measurement method. The objectives of Study 1 include: (a)

confirming the main motivator of online review searching behavior and the type(s) of products/services they commonly search reviews for, and (b) understanding respondents search behaviors in online review websites, including how they sample or proceed with reading the reviews and how they utilize various search functions to facilitate their search process.

The study followed a 2 (average rating level: high vs. low) $\times$ 2 (product rating distribution: present vs. absent) design. The average rating level is a between-subject factor and the rating distribution presence is a within-subject factor. The participants were given a prepurchase scenario when approaching the review websites (rating distribution presence vs. absent) one by one. Two sample websites were shown to each participant, namely, tripadvisor.com and booking.com. Tripadvisor.com represents the distribution presence condition and the booking.com represents the distribution absence condition. Both tripadvisor.com and booking.com have a high average rating version and a low average rating version. 29 respondents (age 18-35) with prior experience in searching online reviews participated in Study. 17 respondents are under the high average rating condition while 12 are under the low average rating condition (randomly assigned).

At the beginning of the Study 1, the interviewer asked questions about the main motivators of online review searching behavior and the types of products/services they commonly search reviews for. Subsequently, the

interviewer showed the participants a sample online review discussion from popular review websites (e.g., tripadvisor.com) and asked the participants to think aloud as they performed a search task on the review websites. They were asked to say whatever they were thinking, doing, and feeling, as they performed their task. In order to understand how the participants sample or proceed in reading the reviews, and how they utilize various functions to facilitate their search process, the interviewer would remind the participants to think aloud as they were viewing and ask them to explain/elaborate points they made that were most relevant to the current study. A think-aloud protocol guide was designed (see Appendix A) to explore the views and opinions of the respondents and to prevent the possible omission of any important information (Patton 1990).

The eye movements were captured using an eye-tracker to provide more understanding of the viewing process. Applications of eye-tracking have appeared in substantive areas of visual marketing (Pieters and Warlop 1999; Russo and Leclerc 1994). Utilizing the eye tracking method allows the researcher to examine the exact area (e.g., a particular review message) that each consumer is paying attention to and the amount of attention paid to the area within a page. Other process-tracing methods, such as clickstream tracking, are not able to differentiate how much attention a consumer pays to each area within a page when no clicking or scrolling is involved. In addition, the eye-tracking method has the advantage of being far less obtrusive than alternative process-tracing methodologies such as

verbal protocols (Russo 1978). When applied to the study of online review search behavior, eye-fixation data can provide precise details in within-page information search and processing behavior, including the exact message that each consumer is processing to and the amount of effort involved. Thus, recording of consumers' eye movement patterns is very well suited to study spontaneous search behaviors.

Eye movements, including fixations and saccades, are a necessary and integral part of information acquisition (Rayner 1998, 2009; Russo 1978). During fixations (moments when the eye remains relatively stationary), information is extracted from the stimulus upon which the eyes fixate, and processed internally. During saccades (rapid movements between fixations), vision is suppressed (Rayner 1998, 2009; Russo 1978). Two frequently used measures of information processed in marketing research are fixation frequency and gaze duration (Aribarg, Pieters, and Wedel 2010; Pieters 2008). Fixation frequency refers to the total number of fixations. Gaze duration represents the sum of all consecutive fixation durations made on a region, such as a review message. Prior studies suggest that gaze duration reflects the amount of information acquired from the stimulus and the extent to which the information is processed (Lam, Chau, and Wong 2007; Rayner 1998; Wedel and Pieters 2000). In addition, previous research shows that fixation frequency and gaze duration is highly correlated (Russo 2011). Thus, the current search measurements focus on gaze durations only. The reviews provided are categorized into two sets: (a) reviews that are consistent with the average rating

(consistent reviews); and (b) reviews that inconsistent with the average rating (deviant reviews). In the current research, the unit of analysis is a particular set of reviews instead of each word. The total search durations within each set of reviews are captured using eye-tracking method and compared in the results section.

Operationalization of the Review Classification

The reviews under a threaded discussion are categorized into two sets based on their corresponding product ratings and the average product rating: reviews that are consistent with the average rating (consistent reviews) and reviews that deviate from the average (deviant reviews). A relatively broad definition of consistent reviews is adopted. That is, the reviews were categorized into consistent reviews and deviant reviews based on the valence. If the average rating is 7.3 out of 10, reviews with ratings 6 to 10 were classified as the “consistent reviews”; others were the “deviant reviews”. If the average rating is 3.7 out of 10, reviews with ratings 1 to 5 were considered the “consistent reviews”; others were the “deviant reviews”. This classification is justified on the basis of prior work on online reviews and the empirical evidence that I gathered. My justifications are elaborated as follows.

The classification adopted is similar to the method proposed by Chen, Liu and Zhang (2011) whereby the reviews are categorized into three types based on their relative valence: relatively negative, relatively positive or relatively neutral. The consistent reviews in the current research correspond to the “relatively neutral”

review category defined by Chen, Liu and Zhang (2011) and the deviant reviews correspond to a combination of both “relatively negative” and “relatively positive” review categories.

The “relatively positive” and “relatively negative” reviews are combined into one category for the following reasons. First, the key relationships under investigation in the current research are different from Chen, Liu and Zhang’s work (2011). The novel element of the current research framework is to explain and predict how the presence of certain search aids (e.g., an average rating or a rating distribution) influences the search of reviews in the search stage by extending the anchoring literature. Thus, the empirical studies should focus on this element of the framework. The classification between the consistent and deviant reviews sufficiently enables the testing of the anchoring effect of average rating presence, and how the provision of a dispersed distribution influences the consumers’ search behaviors by “forewarning” about the “anchor”. If such anchoring effect is only tested under a single high or low average rating setting, it would be difficult to rule out the alternative explanation that the impact was driven by the valence of the reviews instead of the consistent/deviant review classification. To evade the valence-based alternative explanation in the search stage, the effects of the search aids are tested under both a positive and a negative average rating condition. The current classification of consistent and deviant reviews is appropriate for tests of the proposed model.

Second, consumers may consider a number of reviews with ratings that are “close” to or consistent with the average rating as the consistent reviews or the “relative neutral reviews”. Such definition of the consistent reviews could be developed based on a relatively narrow or broad range within the 1 to 10 rating scale. For example, following a relatively narrow definition, with an average rating of 7.3 out of 10, reviews with ratings 7 or 8 could be classified as the “consistent reviews”; others are the “deviant reviews”. On the other hand, following a relatively broad definition, with an average rating of 7.3 out of 10, all the reviews with ratings 6 to 9 or even 6 to 10 could be considered as the “consistent reviews”. The latter constitutes the valence-based classification of consistent reviews (the current scheme). Based on the qualitative results of study 2, consumers view consistent reviews or deviant reviews in a very broad sense. Most of them simply classify the reviews as “positive” or “negative” and consider all positive reviews as “consistent” with a high average rating. Furthermore, consumers who are not given the average rating only have a vague idea about where the average rating lies. Therefore, a broader range for classifying the consistent reviews is justified. In addition, the quantitative results on the structural model in Study 2 also support such categorization. The data analysis based on the two alternative narrower classifications of reviews did not show the expected results. Most of the proposed relationships were insignificant and the SEM model fit was poor (see Appendix D). When the reviews are classified based on the broad definition (current scheme),

many of the predicted relationships are supported and most of the fit indices become acceptable. This indicates that the broad classification is a more realistic representation of the data.

Third, in the empirical studies, only participants that inspect the full content of reviews from all types of reviews are included in the usable sample. This is because of the decomposition of the search variables. A full view ratio is the quotient of the number of reviews read divided by the total number of reviews previewed within a particular type of reviews. However, for respondents that did not preview a particular type of reviews, their total numbers of reviews previewed within consistent reviews equal to zero. The full view ratios are not applicable when the denominators are equal to zero for these cases; thus their full view ratio are indicated as systematic missing values. Furthermore, average search depth is the quotient of the total detailed view search time divided by the number of reviews read within a particular type of reviews. Similarly, for respondents that did not read a particular type of reviews in detail, their total number of reviews read within a particular type of reviews equal to zero and the average search depth is not applicable as the denominators are zeros; thus their average search depth is indicated as systematic missing values. In SEM, responses with systematic missing values are removed list-wise and only responses without any missing values are used for analysis. The sample size for Study 2 is 312 (after removing outliers). Under the broad categorization (current scheme), there are a total of 247

usable responses and a total of 94 these cases with systematic missing values (those cases described above). If the deviant reviews are further divided into positively and negatively deviant reviews, according to the decomposition of search variables, only responses that read (clicking on “read more”) at least one consistent review, one positively deviant review and one negatively deviant reviews are included in the usable sample. Many participants only inspect the full view sections of positively or negatively deviant reviews. If the deviant reviews are further divided into positive and negative categories, the usable sample size would be greatly reduced. Besides, the current broad categorization also helps increase the usable sample size. Under the two alternative narrower classifications, only 207 and 244 usable responses were included in the data analysis respectively after removing the systematic missing values caused by the respondents not reading the full views of either consistent or deviant reviews. As both alternative classifications of the consistent reviews cover comparatively narrower ranges of reviews, the chances that participants do not read the full views of such relatively narrowly defined consistent reviews are higher.

Lastly, in Study 2, 3.7/10 and 7.3/10 are chosen to represent a low and a high average rating respectively. Based on the broad definition of consistent reviews, only negatively deviant reviews are available for a discussion with a positive average rating while only positively deviant reviews are available for a discussion with a negative average rating. Thus, the construct “deviant reviews” is used to

generalize both conditions. Such broad definition of consistent reviews is also applicable for a neutral average rating (e.g., 5.5/10). In this case, it is acknowledged that a further division of positively and negatively deviant reviews is necessary in order to investigate the impact of reviews in the evaluation stage.

A demographics form was also prepared for the respondents to fill in. The form captures data like age, race, occupation and educational level. Respondents were recruited in a major university in Singapore. The think-aloud protocols are conducted face-to-face in the CSI lab in Nanyang Business School. To conform to the ethical standards, the protocols of both Studies 1 and 2 were approved by NTU-IRB. The consents of the participants were sought prior to the study (McCracken 1988). They were informed about the objective of the study and assured of anonymity. Each session lasted about 30 minutes, was audio-recorded and subsequently transcribed verbatim. As appreciation for their generosity with their time, each respondent was given \$6. The transcripts and audio files were then given to another independent coder to verify the accuracy of the transcription before being used for analysis.

Summary of Results of Study 1

Data Coding

Coding categories were developed based on the assumptions and the data patterns. First, some initial coding categories were generated from the assumptions/predictions based on the literature review (previous research findings).

These predictions include the main motivator of online review searching behavior (prepurchase v.s. knowledge building), the type(s) of products/services they commonly search reviews for (e.g., hotel and travel destinations) and the types of reviews to search (e.g., searching the reviews sequentially). For example, based on prior research on online review search motives (Goldsmith and Horowitz 2006; Hennig-Thurau and Walsh 2003), consumers are expected to search online reviews in order to make an optimal purchase decision or to learn more about a product or product category. Thus, two categories such as “prepurchase search motive” and “knowledge building motive” are developed. In addition, past research has established that travel planning decisions are sensitive to WOM (Freedman 2008). Thus it is expected that hotels/travel destinations could be a popular topic in online review forums and a category named “searching reviews on hotels/travel destinations” is developed. Lastly, as proposed in the baseline model in the absence of any search aids, consumers are expected to search review sequentially. Therefore, a category named “searching the reviews sequentially” is developed.

Secondly, additional categories were generated based on the data patterns. These categories were generated following the procedures suggested by Bogdan and Biklin (1998). Initial coding was conducted by generating numerous category codes as reading the responses, labeling data that was related without worrying about the variety of categories. Then, focused coding was used to eliminate and combine repeating ideas. These additional categories derived from the data include

the type(s) of products/services they commonly search reviews for (e.g., IT products and restaurants), the uses of average ratings and distributions, and reviews to search (e.g., any random reviews, the positive reviews, the negative reviews or both). For example, when a respondent mentioned “I search reviews for IT products” in her response, a new category named “searching reviews on IT product” is developed base on the data.

Two independent coders linked all specific quotes from each response to identified categories. Each column represents a category and each row represents a participant. If a respondent mentioned a particular category in her statements (e.g., “I search online reviews when I want to buy something”), a “1” was assigned under the category (e.g., “prepurchase search motive”) in her row. Otherwise, if she did not mention anything that could relate to the category in her response, a “0” was assigned under that category in her row. After all the responses are coded, the proportional reduction in loss (PRL) is used to assess the measurement reliability (Rust and Cooil 1994). As there are 4879 pairwise agreements out of a total of 5307 possible, the proportion of agreement is .919. The corresponding PRL reliability is .92 for two categories and two judges, which is greater than the minimum standard of .70 for exploratory work (Rust and Cooil 1994). Thus, the coding is considered reliable.

Summary of Qualitative Findings

The qualitative findings are presented under two distinctive “themes”: (a) descriptions about the search process on the two websites, including the search aids that the participants paid attention to and used, and the type(s) of reviews they searched for; and (b) descriptions about their general everyday review search behaviors.

General Online Review Search Behaviors

This exploratory study confirmed many of the assumptions and predictions. First of all, 72.4% (21 out of 29) of the participants indicated they search reviews for prepurchase purpose (most popular motive). The second commonest motive is building product knowledge (24.1% or 7 out of 29 participants). Secondly, 69% of the participants (20 out of 29) mentioned hotel or travel destinations as the target product or service when they searched for reviews online. Other than hotels, IT products (75.9% or 22 out of 29) and restaurants (51.7% or 15 out of 29) were also cited by many participants as the target product or service. Lastly, 75.9% of the participants (22 out of 29) confirmed that they would consider going to a beach resort in Southeast Asia for holidays. Table 6 tabulates the frequencies of the general online review search behaviours.

Table 6
Qualitative Results on General Online Review Search Behaviors

		Frequency	Percentage
Search motives	Prepurchase	21(out of 29)	72.4%
	Building product knowledge	7 (out of 29)	24.1%
Target product	IT products	22 (out of 29)	75.9%
	Hotel or travel destinations	20 (out of 29)	69.0%
	Restaurants	15 (out of 29)	51.7%
	Considered going to a beach resort in SEA for holidays	22 (out of 29)	75.9%
	Search aids		
	Found average ratings useful	22 (out of 29)	75.9%
	Found product rating distributions useful	18 (out of 29)	62.1%
	Used sorting functions	10 (out of 29)	34.5%
Uses of average ratings	Building the first impression	14 (out of 22)	63.6%
	The basis for judgment of consensus	4 (out of 22)	18.2%
	Uses of distributions		
	The basis for judgment of consensus	10 (out of 18)	55.6%
	Building the first impression	5 (out of 18)	27.8%
	Monitoring the number of extreme reviews	4 (out of 18)	22.2%
	Comparing the number of positive and negative reviews	3 (out of 18)	16.7%
Reviews to search	Approaching the reviews sequentially	17 (out of 29)	58.6%
	Reading both positive and negative reviews	11 (out of 29)	37.9%
	Selecting random reviews to read	6 (out of 29)	20.7%
	Constantly looking out for negative reviews	6 (out of 29)	20.7%
	Constantly looking out for positive reviews	5 (out of 29)	17.2%

Regarding participants' responses on general everyday experience of searching online reviews, 75.9% (22 out of 29) of the participants mentioned that they found average ratings useful and 62.1% (18 out of 29) indicated that they found product rating distributions useful. However, only 34.5% (10 out of 29) used sorting functions. Thus, the subsequent study (Study 2) focuses mainly on the first two search aids, as they are more often used in the review search process. Among those who found average ratings useful, the most popular uses of average ratings included building a first impression (63.6% or 14 out of 22 participants mentioned it) and using it as the basis for judgment of general consensus (18.2% or 4 out of 22 participants mentioned it). Among the participants that mentioned the distribution as a useful tool, 55.6% (10 out of 18) of them used it as the basis of judgment for consensus, 27.8% (5 out of 18) of them used it to build a first impression, 22.2% (4 out of 18) of them used it to monitor the number of extreme reviews and 16.7% (3 out of 18) of them used it to compare the number of positive and negative reviews (See Table 7). For those who found search aids not useful, one participant mentioned that she would "read the review contents rather than choosing reviews based on the ratings"; and another mentioned that she would "form her own judgment of the distribution". Also, one participant mentioned that she usually did not notice such distributions.

In terms of the search process, many participants would approach the reviews sequentially (58.6% or 17 out of 29 participants mentioned it). In addition, 37.9%

(11 out of 29) would read both positive and negative reviews, “so that I can get a balanced view”. Furthermore, 20.7% of the participants (6 out of 29) would select random reviews to read. Some participants (20.7% or 6 out of 29) would constantly look out for negative reviews, “so that I know the worst scenarios”. However, some would constantly look out for positive reviews (17.2% or 5 out of 29). These participants commented that “if I keep looking at the bad ones, I would never be able to make a decision”. Two participants mentioned that they would search the deviant reviews and two participants indicated that they would search the moderate reviews. The former two participants indicated that they would “constantly look out for reviews that are inconsistent with the average to get a balanced view”. The latter two found moderate reviews “more persuasive” than the extreme ones.

Search Process on the Two Websites

The frequencies of their indicated search patterns are tabulated across different conditions in Tables 7 and 8. Pearson χ^2 statistic was used to test the two-way associations with adequate frequencies in the cells. According to Tabachnick and Fidell (2007), to examine two-way associations, all cell expected frequencies need to be more than one, and no more than 20% of all cell expected frequencies are less than five. Thus, Pearson χ^2 tests were not conducted on the

categories with cells that contain frequency counts that were below five. In Tables 7 and 8, none of the Pearson χ^2 tests was significant at .1 level.

Table 7
Frequency Distribution of Respondents According to Uses of Search Aids on the Two Websites

For Participants Who Used Average Rating		Average Rating	Distribution		
		Level	Present	Absent	Total
		High	10	11	21
		Low	4	7	11
		Total	14	18	32
Why?	First Impression	Average Rating	Distribution		
		Level	Present	Absent	Total
		High	6	7	13
		Low	3	7	10
		Total	9	14	23
Why not?	Prefer to inspect and choose reviews	Average Rating	Distribution		
		Level	Present	Absent	Total
		High	0	2	2
		Low	1	2	3
		Total	1	4	5
	The average rating does not match the reviews	Average Rating	Distribution		
		Level	Present	Absent	Total
		High	2	1	3
		Low	4	1	5
		Total	6	2	8
For Participants Who Used Distribution		Average Rating Level	Total		
		High	Low		
		9	7	16	
Why?	Forming a judgment on the majority opinion	Average Rating Level	Total		
		High	Low		
		6	5	11	
	Comparing the number of positive and negative reviews	Average Rating Level	Total		
High		Low			
3		2	5		
Why not?	Not noticing the distribution	Average Rating Level	Total		
		High	Low		
		5	1	6	

Table 8
Frequency Distribution of Respondents According to Reviews Searched on the Two Websites

Reviews to Search		Distribution		Total
<i>Negative reviews</i>	Level	Present	Absent	
	High	8	7	15
	Low	3	3	6
	Total	11	10	21
		Distribution		Total
<i>Both positive and negative reviews</i>	Level	Present	Absent	
	High	4	7	11
	Low	3	4	7
	Total	7	11	18
		Distribution		Total
<i>Positive reviews</i>	Level	Present	Absent	
	High	1	0	1
	Low	2	2	4
	Total	3	2	5
		Distribution		Total
<i>Moderate reviews</i>	Level	Present	Absent	
	High	1	1	2
	Low	2	1	3
	Total	3	2	5
		Distribution		Total
<i>Deviant reviews</i>	Level	Present	Absent	
	High	2	2	4
	Low	0	2	2
	Total	2	4	6
		Distribution		Total
<i>Latest reviews</i>	Level	Present	Absent	
	High	4	4	8
	Low	4	4	8
	Total	8	8	16

Referring to “For Participants Who Used Average Rating” in Table 8, when the distribution was absent, 18 participants used the average rating in the search process. When the distribution was present, 14 participants used the average rating in the search process. Although the chi-square test between the uses of average rating and the distribution presence was not significant ($\chi^2(1) = 1.115, p = .291$), the results provide some directional support to the notion that the distribution presence diverts people’s attention away from the average rating. In addition, the results show that when the distribution was absent, more participants used the average rating to build a first impression (14 participants), as compared to those under the condition that the distribution was present (nine participants; $\chi^2(1) = 1.801, p = .180$; refer to “Why? First impression” caption in Table 8). The difference was mainly contributed by the participants under the condition that the average rating is low. Specifically, when the average rating is low, three participants that are under the distribution present condition and seven participants under the distribution absent condition used the average rating to build a first impression ($\chi^2(1) = 2.743, p = .098$). When the average rating is high, six participants that are under the distribution present condition and seven participants under the distribution absent condition used it to build a first impression. This is consistent with the prediction on forewarning effect in the anchoring literature (Epley and Gilovich 2005; Tversky and Kahneman 1974). Other uses of average

rating consist of comparing the average ratings of the same product across different reviews sites (two participants) or with other products (one participant).

For those who did not use the average rating, five participants mentioned that they preferred to inspect and select the reviews based on review contents rather than the ratings and eight participants indicated that they felt that average rating did not reflect the totality of all the reviews. For example, one commented that “the average rating is high. However, I saw some negative reviews. Thus, I think the average rating and the actual review scores does not match”.

Nine participants in the high average rating condition and seven participants in the low average rating condition used distribution. The major uses of distribution include forming a judgment on the majority opinion (11 participants mentioned it) and comparing the number of positive and negative reviews (five participants mentioned it; refer to “For Participants Who Used Distribution” in Table 8). Major reasons that one did not use the distribution include not noticing the distribution (six participants mentioned it) and not able to understand the distribution (two participants mentioned it; not included in Table 8 due to low frequency). Among those who did not notice the distribution, five were under the high average rating condition and only one was under the low average rating condition. This shows that negative majority ratings revealed by the distribution is less common and more diagnostic and therefore attracts more attention.

When the average rating was low, participants were more likely to search positive reviews. All participants that searched other types of reviews distributed evenly among the different conditions (see Table 8).

Eye Movement Data Analysis

Search time spent on consistent reviews was compared with time spent on deviant reviews. When the average rating is high, reviews with ratings 6 to 10 are the consistent reviews and the remaining reviews are the deviant reviews. When the average rating is low, reviews with ratings 1 to 5 are the consistent reviews and the remaining reviews are the deviant reviews. The average rating was present across all conditions. Search time spent on the consistent and deviant reviews under the different conditions, average rating levels (high vs. low) and product rating distribution (present vs. absent), were also compared using independent sample T-test. The results are summarized in the next section.

Results show that participants spent significantly more time on consistent reviews than deviant reviews ($t(110)=0.066$, $p=.000$, $M_{consistent} =49.151$ and $M_{deviant} =13.1558$). This is consistent with the anchoring effects in the model. As the average rating was present across all conditions, people should spend more time on consistent reviews.

The presence of distribution has a significant negative effect on search time spent on deviant reviews ($t(54) = -2.835$, $p=.006$; $M_{present} =6.242$, and $M_{absent} =20.069$). A rating distribution is often used as a judgment basis for the majority

opinion. The stimulus chosen in Study 1 was a unimodal distribution with some dispersion, in which the mode is largely consistent with the average rating. For a unimodal distribution, the mode is very prominent, as there is only one peak. Thus, consumers may focus on the mode rather than the extremes. Thus, when it was present, participants' attention was diverted away from the deviant reviews.

The presence of distribution also had a directional negative effect on search time spent on consistent reviews ($t(54) = -.545$, $p=.588$; $M_{present} = 46.225$, and $M_{absent} = 52.077$). Although the effect was not significant, the direction was consistent with the forewarning effect in anchoring literature. No significant differences were found between the high and low average rating groups in terms of their search time spent on both consistent reviews ($t(54) = 1.207$, $p=.209$; $M_{high} = 54.991$, and $M_{low} = 41.364$) and deviant reviews ($t(54) = -.994$, $p=.325$, $M_{high} = 10.927$, and $M_{low} = 16.128$).

Conclusion

The qualitative results of Study 1 confirm prepurchase motive as the main motivator of online review searching behavior and the relevance of the target product/service—a beach resort in Southeast Asia. The results also confirm the relevance of the two search aids, namely, the average rating and rating distribution. Furthermore, the results provide some support to the prediction that the distribution presence diverts people's attention away from the average rating. In addition,

results show that participants were more likely to use the average rating to build a first impression when the distribution was absent.

The quantitative results of Study 1 show some preliminary support to the anchoring effect of the average rating presence. Some preliminary evidence was found on the forewarning effect of the distribution presence on the consistent reviews.

CHAPTER 5

STUDY 2

Study 2 consists of three components: a pretest, a main study (eye-tracking experiments) and interviews. The purpose of the pretest is to refine the experimental stimuli and measures designed for the main study. Then, the main study (eye-tracking experiments) was conducted to test the hypothesized relationships. Lastly, through the interviews, the study aims to explore and uncover other search strategies and/or considerations that people have during the search process in order to complement the proposed model.

Methods

Experimental Design

The objective of study 2 is to investigate how various forum search aids (provision of an average rating and provision of a distribution of product ratings) influence review searching behaviors and ultimately the corresponding search outcomes. In this study, I give the participants a prepurchase scenario when approaching the review website. To create a fair test of the conceptual models, a target product/service that is responsive to online reviews is needed. Since past research has established that travel planning decisions are sensitive to WOM (Freedman 2008) and hotels/resorts are a very popular topic in online review forums nowadays based on Study 1 results, a holiday beach resort is selected as the target product in the study.

The research employs a 2 (average rating level: high vs. low) $\times$ 2 (average product rating: present vs. absent) $\times$ 2 (product rating distribution: present vs. absent) full factorial between subjects design. Eight versions of threaded review discussions (websites) based on the above eight conditions are created as the experimental material (see Appendix B for sample websites on the average rating level manipulation). The manipulation of the average rating level is based on the actual average rating of the discussion - the mean of all the review ratings in the threaded discussion. In the condition that the average rating is high, the discussion contains a majority of reviews with high ratings. In the condition that the average rating is low, the discussion contains a majority of reviews with low ratings. The average rating presence manipulation involved varying the design of the websites by including or excluding the average rating statistic from the top of the threaded discussions. Similarly, the rating distribution presence manipulation also involved varying the design of the websites by including or excluding the product rating distribution from the top of the threaded discussions. Review helpfulness voting/sorting function is not available across all conditions. Sorting by rating function is present across all conditions and participants can choose to use or not to use it. Both eye movements and structured questionnaires were used as measurement tools. Demographic information was also measured at the end of the questionnaires. Foreigners (non-PRs) are excluded from the sample because of the language experience issue.

Beside forum designs, certain personal characteristics such as the need for closure (NFC) (Webster and Kruglanski 1994) and need for cognition (NC) can affect the information processing of individuals (Cacioppo et al. 1996; Verplanken, Hazenberg, and Palenwen 1992). According to Webster and Kruglanski (1994), NFC leads to "seizing" and "freezing" upon accessible information at the early stage and insensitivity to subsequent inconsistent information. Literature on NFC emphasizes that people high in overall NFC have a greater tendency to use heuristic cues, such as consensus (Kruglanski et al. 2006), or numerical anchors (Kruglanski and Freund 1983) in order to reach urgent judgments/closure. In the current context, this suggests that NFC may influence the search behaviors. Need for closure is measured by the revised scale (Roets and Hiel 2011). Besides, prior research also suggests that individual differences in NC can affect the information processing of individuals (Cacioppo et al. 1996; Verplanken et al. 1992). NC refers to "an individual's tendency to engage in and enjoy effortful cognitive endeavors". Individuals with high NC have a tendency to acquire, organize, evaluate, and synthesize information (Cacioppo et al. 1996). They tend to select more information and exert more task-related cognitive effort (Verplanken et al. 1992). Therefore, both NFC and NC are included as covariates during the information search stage.

In addition, research also suggests that the effectiveness of word-of-mouth information can be influenced by its consistency with the prior attitudes towards

the product (Cheung et al. 2009; Zhang and Watts 2008). Online word-of-mouth information that confirms prior beliefs is found to be more influential than that disconfirms them. Therefore, in the current research context, the absolute difference between prior attitude and the average rating is included as a covariate to control for its effect during both the search stage and product evaluation stage. Other control variables include prior product knowledge, product involvement and prior online review experience, as they potentially affect consumers' search behavior (Bettman and Park 1980; Park et al. 2007; Ratchford 2001).

Measurements of Search Behaviors

Similar to Study 1, the search behaviors are measured by eye-tracking method. However, not only the total search durations within each set of reviews are captured, but the number of reviews previewed and read (clicking on “read more”) are also measured. Subsequently, the preview ratios and the average search depth are derived from the total search time, number of reviews previewed and read within each set of reviews. Table 9 summarizes the definitions and measures of the search behavior constructs.

Table 9
Definitions and Measures of the Search Behavior Constructs

Construct	Definition	Measure
<i>Number of Reviews Available</i>	Number of reviews present within each set on the website	Total number of review message shown within each set on the website
<i>Preview Ratio</i>	Within each set of reviews, provided that a review is present, the probability of this review being previewed	Within each set of reviews, the total number of review message previews that at least one eye-fixation land in divided by the numbers of reviews available
<i>Number of Reviews Previewed</i>	Number of reviews previewed within each set	Total number of review message previews that at least one eye-fixation land in within each set of reviews
<i>Full View Ratio</i>	Within each set of reviews, provided that a review is previewed, the probability of this review being read (clicking on “read more”)	Within each set of reviews, the number of reviews read divided by the numbers of reviews previewed
<i>Number of Reviews Read</i>	Number of reviews read within each set (clicking on “read more”)	Total number of review messages that one clicks on “read more” to inspect the details within each set of reviews
<i>Average Search Depth</i>	Level of processing within each set of reviews	Total gaze duration within each set of the reviews read divided by the number of reviews read
<i>Search Time Allocation</i>	Total amount of time a consumer devotes into reading each set of the reviews	The total gaze duration within each set of the reviews read

Only reviews read (clicking on “read more”) are considered in the search time allocation. This is because the previews in most review forums are very short and usually contain mostly the background information (e.g., the consumption scenarios). Very little information about the focal products/services appears in the

review previews, especially when the reviews are of median to long length (e.g., 7-10 arguments in the current research context). Following the above definition, I derive from the eye-fixation records measures of the key constructs under examination.

Materials

This experiment puts the subjects in the context of considering an unfamiliar beach resort named Kamala Beach Resort. It is a beach resort located in Phuket, Thailand. The use of unfamiliar and neutrally rated brand names can help to maximize the impact of interpersonal influence such as online reviews and avoid any possible biases and prior impressions the subjects might have (Zemboirain 2007).

Eight versions of threaded review discussions (websites) based on the eight conditions are created as the experimental material. The eight conditions follow a 2 (average rating level: high vs. low) $\times$ 2 (average product rating: present vs. absent) $\times$ 2 (product rating distribution: present vs. absent) full factorial between subjects design. One of the most popular hotel review website, Tripadvisor.com, has an average of about 300 reviews displayed for each hotel (Ross 2010). Additionally, on Tripadvisor.com, a sample of hotels in several major cities in south-eastern Asia has an average about 230 reviews. Therefore, I included 181 reviews in the simulated threaded discussion. Owing to the focus of the current research, I assume that the average product ratings and product rating distributions

do not vary over time (counter-balanced). That is, reviews from each time period follow a similar distribution as the overall reviews. The 181 reviews are arranged in preview formats and divided into 19 pages and participants are free to click on the “read more” button to inspect the full messages or “read less” to go back to the previews. The corresponding product rating is presented at the top of each review. On the first page, the resort name, the summary statistics and the sorting button are presented before the reviews. Half of the participants were provided with a product average rating, which varies across two conditions: positive and negative. Half of the participants were provided with a product rating distribution, while the others were not. The sorting by rating button was present across all different groups. All reviews were first selected from existing review websites (e.g., tripadvisor.com). Three independent raters rated the reviews on the word count, number of arguments, message content-rating consistency, and message strength of each review independently and the inter-rater reliabilities are satisfying. All Cronbach's α and PRL reliabilities among 3 raters are greater than the minimum standard of .70 for exploratory work (Rust and Cooil 1994). Based on the agreement of at least two raters, only strong and consistent messages with median length 148-223 words (including title) and 7-10 unique arguments were included in the experimental materials. Such word counts and numbers of arguments represent the typical length of a consumer review on popular forums. No significant differences in word

counts, argument counts, message content-rating consistency and message strength are found among reviews of different ratings using ANOVA tests.

Procedure

A group of students are recruited from a major university in Singapore through an online external experimental research sign-up system. All have normal or corrected-to-normal vision and have an adequate level of English language proficiency. These respondents participated in exchange for \$6. Fifteen of them joined the pretest of Study 2 to test the experimental procedure, questionnaires and usage of the eye trackers. The sample size for Study 2 is 312 (after removing outliers) and there are a total of 247 usable responses without any missing values. The removal of outliers and missing values is covered in the results section. The demographic profile of the 247 usable responses was as follows. The sample consisted of 90 male (36.4%) and 157 female (63.6%) respondents. 91.5% of the respondents fell between the age of 19-24, and 8.5% of the respondents were above 24 years old. The median age of the respondents was 22. 98.4% of the respondents were undergraduate students, and the rest were graduate students.

Participants in the experiment were randomly assigned to one of the eight conditions in a laboratory setting. Each participant was seated at a computer, which contains the instructions, the site and the survey (See Appendix B for the protocol). This experiment put the participants in the context of considering an unfamiliar beach resort named Kamala Beach Resort. The eye tracker was first

calibrated, after which the participants read the instructions presented on the computer screen analogously to the experimental website. Participants were told that they would be visiting a portion of a review web site devoted to a specific beach resort named Kamala Beach Resort. They were then told that, “you have heard of the above beach resort from advertisements and you decide to search more about the resort in online reviews”. They were first given a general description of Kamala Beach Resort from the advertisements as follows:

‘Kamala Beach Resort is one of the all-Inclusive resorts on the Phuket, Thailand. The resort has a variety of rooms, suites, cottages, casual and elegant restaurants, bars, a full service spa and water activities to enjoy. The resort is 35 minutes drive to Phuket Airport, 15 minutes drive to Patong, and 10 minutes walk to Kamala Village and Phuket Fantasea. Guest room rates range from SGD 70-150, depending on the room type and season’.

A sample of the advertisement is included in Appendix B - Study 2 Materials. Then, they were then asked to rate their attitude towards Kamala Beach Resort, intention to visit the resort as their travel destination and their evaluation confidence on a seven-point scale. These ratings would form the baseline against which the post-search ratings are compared. The rating task was followed by a filler task. After the filler task, they were presented with a threaded review

discussion on a beach resort on a simulated review discussion. Focal manipulations and some peripheral information about the resort were included in the scenario. All participants were instructed to imagine that they were considering Kamala Beach Resort as their travel destination in the coming school break in two months time. Before doing so, they should type two questions they had about this resort. To ensure that they would be attending to information relevant to them, I did not specify what to look for. These instructions are used in prior research to instill a prepurchase search goal (Schlosser 2003; Schlosser, White, and Lloyd 2006).

Participants then visited the site. Their eye movements over the section were tracked by using Tobii T120 eye trackers to provide the needed search measures. The Tobii eye trackers use infrared corneal reflection methodology, which measures the distance and angle of the reflection of infrared light from the center of the pupil to determine the point of fixation of the person, after calibration. These devices integrate miniature head- and eye-tracking cameras into an LCD monitor, which enables individuals to freely move their heads during eye-tracking. The online activities were also video-recorded digitally by the PC using software for data coding purpose. At the beginning of the session, a fixation point appears on the screen for calibration and the participant looks at the point and the eye tracker automatically adjusts the calibration if necessary. After viewing the website, the participants filled in the online questionnaire containing the

measurement scales on changes in consumers' attitude and purchase intention towards the focal product, evaluation confidence, control variables (including need for closure, need for cognition, knowledge about resorts, involvement in resorts, prior online review experience, and prior attitude towards the resort), and demographics (see Appendix C). Each test session lasted for approximately 30min.

Qualitative Analysis of Interviews in Study 2

Purpose for Conducting Interviews in Study 2

The objectives of conducting face-to-face interviews in Study 2 are three folds. First, the interviews aim to confirm some of the previous assumptions made in the paper. For example, it was assumed that people search reviews sequentially by the date in the absence of any search aids. The second objective is to validate some of the predictions on search behaviors apart from the quantitative measures, including how they sample or proceed with reading the reviews and how they utilize various search aids to facilitate their search process. Lastly, through the interviews, the study aims to explore and uncover other search strategies and/or considerations that people have during the search process in order to complement the proposed model.

Data Collection

The study follows a 2 (average rating presence: present vs. absent) $\times$ 2 (product rating distribution presence: present vs. absent) $\times$ 2 (average rating level:

high vs. low) between-subject design. Participants were given a prepurchase scenario in the cover story when approaching the review website. After viewing the website, the participants filled in an online questionnaire which contains the quantitative measurement scales on changes in consumers' attitude and purchase intention towards the focal product, evaluation confidence, control variables (need for closure, need for cognition, knowledge about resorts, involvement in resorts, prior online review experience, and prior attitude towards the resort) and demographics. Subsequently, face-to-face interviews were conducted with about 20% of the participants (randomly selected). Questions on how they think and feel before and during the search process and on their search behaviors were asked, audio recorded and transcribed to verbatim. The verbatim transcripts formed the basis for analysis and interpretation. A total of 87 interviews were conducted. The current study focuses mainly on the participants that read at least one consistent review and one deviant review in detail. Thus, cases that did not fulfill such criteria were removed. As a result, there were 68 cases left. These cases distributed evenly across the eight conditions. No significant association was found across the three factors based on Pearson χ^2 tests ($p > .371$). The number of participants within each cell ranges from 6 to 11.

Developing Coding Categories

Coding categories were developed based on some assumptions/predictions and the data patterns. First, some initial coding categories were generated based on

the predictions in the proposed model. As proposed in the baseline model in the absence of any search aids, consumers are expected to search reviews sequentially. Therefore, a category named “searching the reviews sequentially” is developed. In addition, consumers are hypothesized to display different search behaviors on consistent and deviant reviews. Categories like “searched for consistent reviews” and “searched for deviant reviews” were developed.

Secondly, additional categories were generated based on the data patterns. These categories were developed following the procedures suggested by Bogdan and Biklin (1998). The interview transcripts were first read twice during long, undisturbed periods. Next, initial coding was conducted by generating numerous category codes. The coder read the responses and labeled data that was related, without worrying about the variety of categories. Last, focused coding was used to eliminate and combine repeating ideas. Larger themes that connected codes were identified. Repeating ideas are the same ideas expressed by different respondents, while a theme is a larger topic that organizes or connects a group of repeating ideas (Bogdan and Biklin 1998). Categories that were generated from observing the data patterns include “the average rating helps build the initial impression”, “the distribution indicates the majority opinion”, “searching deviant reviews to reduce uncertainty”, “searching majority positive reviews”, “searching majority negative reviews” “searching both positive and negative reviews”, “evaluating (the resort) based on majority opinion”, “evaluating (the resort) through comparing the pros

and cons”, “evaluating (the resort) based on the number of negative reviews”, and “feeling surprised to see bad reviews”. The eight interview questions (see Appendix B) focus on three distinctive “themes”: (a) descriptions about the search process and the reasons behind such search behaviors, including the search aids that participants paid attention to and used, and the type(s) of reviews they searched for; (b) a description about their evaluation process/strategies; and (c) their feelings during the search process.

Linking Specific Quotes to Identified Categories

Two independent coders linked all specific quotes from each response to identified categories. For example, if one response mentioned that she “plans to read the extreme reviews”, then a “1” was assigned under the category “to search extreme reviews”; otherwise, a “0” was assigned. Using the proportional reduction in loss (PRL) reliability measures (Rust and Cooil 1994), as there are 3214 pairwise agreements out of a total of 3290 possible, the proportion of agreement is .977. The corresponding PRL reliability is .98 for 2 categories and 2 judges, which is greater than the minimum standard of .70 for exploratory work (Rust and Cooil 1994). Thus, the coding is considered reliable.

Summary of Results

Responses under the eight different conditions were compared using cross-tabulation and Pearson χ^2 tests. Pearson χ^2 tests were not conducted on the categories with cells that contain frequency counts that were below five

(Tabachnick and Fidell 2007). The frequencies of their indicated search patterns and evaluation strategies are tabulated across different conditions in Tables 10 to 12.

The first section of the findings details the search process and the reasons behind the search behaviors. Next, their evaluation process is described. The last section presents their feelings during the search process.

Search Aids

Table 10 tabulates the frequencies of their indicated uses of the average rating by the average rating level and the distribution presence. Four participants under high average rating condition and six participants under low average rating condition indicated that they were paying attention to the average rating. This provides some directional support for the notion that negative average is more diagnostic and unexpected, thus it attracts more attention. When the distribution was present, seven out of ten participants explicitly mentioned that they were paying attention to the average rating, while only three stated that they were paying attention to the average rating when the distribution was absent. This shows that the presence of a distribution did not divert their attention away from the average rating.

Table 10

Frequency Distribution of Respondents on Uses of the Average Rating

Paying attention to: *Average rating*

Distribution	Level of average rating		Total
	High	Low	
Present	3	4	7
Absent	1	2	3
Total	4	6	10

Search according to: *Average rating*

Distribution	Level of average rating		Total
	High	Low	
Present	5	5	10
Absent	2	0	2
Total	7	5	12

Why? *Initial impression*

Distribution	Level of average rating		Total
	High	Low	
Present	2	5	7
Absent	1	0	1
Total	3	5	8

Referring to the results under the caption “Search according to: Average rating” in Table 10, seven participants under high average rating condition and five participants under low average rating condition indicated that they were searching according to the average rating ($\chi^2 (1) = .090, p = .764$). This difference is not significant, which suggests that a low average rating did not amplify the usage of average rating in search process as proposed. When the distribution was present, ten out of 12 participants were searching the reviews based on the average rating. Only two participants were searching the reviews based on the average

rating, when the distribution was absent. These two participants were under the condition that the average rating was low. This could be explained by the stimulus used in the experiment (a unimodal distribution with some dispersion). Instead of diverting people's attention away from the average rating, it could also confirm the relevance of the average rating, as the mode of the distribution is largely consistent with the average rating. In the experiment, the mode of the distribution and the average rating matches each other. That is, the distribution can help confirm the usefulness of the average rating. As a negative average rating is less common, consumers are very uncertain about the usefulness of the average rating. For example, they might wonder if the majority opinion was still positive and if there were some extremely negative ratings that brought down the average rating. Thus, presence of the distribution is more essential under such context.

Two reasons that the participants searched according to the average rating were mentioned. First, eight participants indicated that it was used to build an initial impression of the resort (refer to the results under the caption "Why? Initial impression" in Table 10). Second, two participants commented that the average rating was a basis for judgment of the majority opinion (not presented in Table 10 due to low frequency). These two reasons are closely related to each other. Interestingly, when the rating distribution was present, seven participants (88%) mentioned that they used the average rating as an initial impression. Only one used it the same way, when the rating distribution was absent. This also shows that the

distribution used in the experiment did not divert people's attention away from the average rating. Instead, it confirmed the relevance of the average rating.

Table 11 tabulates the frequencies of their indicated uses of the distribution by the average rating level and the average rating presence. Six participants pointed out that they were paying attention to the rating distribution. Four of them were under the condition that the average rating was present and two were under the condition that the average rating was absent. This shows that the average rating presence helps direct some attention to the rating distribution. Eight out of 17 participants under the condition that the average rating was present and nine under the condition that the average rating was absent searched according to the distribution ($\chi^2(1) = .020, p = .889$). This difference is not significant at the .10 level. Ten participants used it as a basis for judgment of majority opinion (see Table 11) and two used it to form an initial impression (not presented in Table 11 due to low frequency). One participant commented that he was trying to find out the average rating by looking at the rating distribution's mode, when the average rating was absent. This reveals that when the average rating is absent, the distribution could serve as a judgment basis for the average rating.

Table 11

Frequency Distribution of Respondents on Uses of the Distribution

Paying attention to: *Distribution*

Level of average rating	Average rating		Total
	Present	Absent	
High	2	1	3
Low	2	1	3
Total	4	2	6

Search according to: *Distribution*

Level of average rating	Average rating		Total
	Present	Absent	
High	7	5	12
Low	1	4	5
Total	8	9	17

Why? *Majority opinion*

Level of average rating	Average rating		Total
	Present	Absent	
High	4	3	7
Low	0	3	3
Total	4	6	0

Search Behaviors

Table 12a presents the frequencies of their indicated search behaviors on consistent and deviant reviews by the average rating level and the distribution presence. Fourteen participants mentioned that they searched deviant reviews. Among those who searched deviant reviews, 11 were under the condition that the distribution was present and three were under the condition that the distribution was absent. This is consistent with the prediction that the distribution presence leads to more search on the deviant reviews. Two mentioned that deviant reviews

were unexpected and three mentioned that they searched deviant reviews to reduce uncertainty caused by inconsistent reviews.

Table 12a
Frequency Distribution of Respondents According to Search Behaviors on
Consistent and Deviant Reviews

Reviews to search	Consistent				
		Level of average rating			
		Distribution	High	Low	Total
		Present	4	4	8
		Absent	3	1	4
		Total	7	5	12
	Deviant				
		Level of average rating			
		Distribution	High	Low	Total
		Present	7	4	11
		Absent	3	0	3
		Total	10	4	14
Why deviant?					
To reduce uncertainty caused by inconsistent reviews		Level of average rating			
		Distribution	High	Low	Total
		Present	0	2	2
		Absent	1	0	1
		Total	1	2	3

Table 12b presents the frequencies of their indicated search behaviors on positive reviews, negative reviews and both reviews by the average rating level, average rating presence and the distribution presence. Seven participants indicated that they searched mainly positive reviews. These seven participants distributed evenly among the different conditions. Twenty-six participants indicated that they searched mainly negative reviews and 12 of them were under the high average

rating condition. According to the classification of consistent and deviant reviews, these 12 participants were actually searching mainly deviant reviews. Among these 12 participants, four were under the condition that the average rating was present, while the other eight were under the condition that the average rating was absent. Eight of them were under the condition that the distribution was present, while the other four were under the condition that the distribution was absent. The results showed directional support to the prediction on the negative effect of average rating presence and positive effect of the distribution presence on deviant reviews.

Table 12b
Frequency Distribution of Respondents According to Search Behaviors Based on Review
Valence

Reviews to Search:			
<i>Majority Positive</i>			
	Level of average rating=high		
	Average rating		
Distribution	Present	Absent	Total
Present	2	0	2
Absent	0	1	1
Total	2	1	3
	Level of average rating=low		
	Average rating		
Distribution	Present	Absent	Total
Present	2	0	2
Absent	1	1	2
Total	3	1	4
<i>Positive and Negative</i>			
	Level of average rating=high		
	Average rating		
Distribution	Present	Absent	Total
Present	4	6	10
Absent	2	6	8
Total	6	12	18
	Level of average rating=low		
	Average rating		
Distribution	Present	Absent	Total
Present	1	4	5
Absent	3	4	7
Total	4	8	12
<i>Majority Negative</i>			
	Level of average rating=high		
	Average rating		
Distribution	Present	Absent	Total
Present	3	5	8
Absent	1	3	4
Total	4	8	12
	Level of average rating=low		
	Average rating		
Distribution	Present	Absent	Total
Present	3	4	7
Absent	4	3	7
Total	7	7	14

Thirty participants indicated they searched both positive and negative reviews in order to get a balanced view (refer to caption “positive and negative” in Table 12b). Among these 30 participants, ten were from the group that the average rating was present and 20 were from the group that the average rating was absent ($\chi^2 (1) = 4.963, p = .026$). This confirms the prediction that when an average rating is provided, people are much more focused during the search process. Thus, they do not search positive and negative reviews in a balanced manner. Search aids direct people’s attention to certain types of reviews.

Table 12c presents the frequencies of sequential review search behaviors by the average rating level, average rating presence and the distribution presence. Twelve participants indicated that they did not search a particular set of reviews, but searched reviews sequentially (sorted by latest date first by default). These 12 participants distributed evenly among the different conditions. The two participants that searched latest and oldest reviews were under the condition that the search aids were absent. All three participants that searched the reviews randomly were under the condition that the average rating was absent (not included in Table 12c due to low frequency). To some extent, this provides some support for the assumption that in the absence of search aids, people search reviews by the date or randomly.

Table 12c
Frequency Distribution of Respondents According to Search Behaviors Based on
Date of Post

Reviews to Search:			
<i>Sequentially(latest first)</i>			
Level of average rating=high			
	Average rating		
Distribution	Present	Absent	Total
Present	4	2	6
Absent	2	3	5
Total	6	5	11

Level of average rating=low			
	Average rating		
Distribution	Present	Absent	Total
Present	2	0	2
Absent	2	1	3
Total	4	1	5

Evaluation Strategies

Table 13 shows the frequencies of their indicated uses of search strategies and their feelings during the search process by the average rating level, the average rating presence and the distribution presence. The ways how participants evaluated the resort using the information gathered were classified under three categories: First, 15 participants were following majority opinion. Second, nine were comparing the pros and cons. Lastly, six were making a product judgment by the number of negative reviews.

Table 13

Frequency Distribution of Respondents on Evaluation Process				
Evaluation Strategies <i>Following majority opinion</i>	Level of average rating=high			
	Average Rating			
	Distribution	Present	Absent	Total
	Present	3	1	4
	Absent	1	4	5
	Total	4	5	9
	Level of average rating=low			
	Average Rating			
	Distribution	Present	Absent	Total
	Present	0	3	3
	Absent	1	2	3
	Total	1	5	6
<i>Comparing the pros and cons</i>	Level of average rating=high			
	Average Rating			
	Distribution	Present	Absent	Total
	Present	1	2	3
	Absent	1	2	3
	Total	2	4	6
	Level of average rating=low			
	Average Rating			
	Distribution	Present	Absent	Total
	Present	0	1	1
	Absent	1	1	2
	Total	1	2	3
<i>Number of negative reviews</i>	Level of average rating			
	Distribution	High	Low	Total
	Present	1	1	2
	Absent	3	1	4
	Total	4	2	6

When the average rating was low and the average rating was present, participants were least likely to follow the majority opinion. Only one participant under this condition chose this strategy. This provides some support to the notion that when the average rating is present, participants are more likely to focus on the average rating and neglect the dispersion of the opinions. A low average rating can strengthen this effect. All the six participants making a product judgment by the number of negative reviews were under the condition that the average rating was absent. Following this strategy, regardless of the number of the positive reviews, if they see many negative reviews, they would form a low evaluation towards the product. No other significant relationship was found between an evaluation strategy and the condition that participant was in (see Table 13).

Feelings

Five participants were surprised to see negative reviews (see Table 14) and two were surprised about the amount of negative reviews (not present in Table 14 due to low frequency). All were under the condition that the distribution was present. Out of the seven participants, two were under the condition that the average rating was high. The other five were under the condition that the average was low. One participant mentioned that she was surprised to see a negative average rating (not present in Table 14 due to low frequency).

Table 14
Frequency Distribution of Respondents on Feelings During the Search Process

Feelings

Surprised to see bad reviews Level of average rating=high

Distribution	Average Rating		
	Present	Absent	Total
Present	1	0	1
Absent	0	0	0
Total	1	0	1

Level of average rating=low

Distribution	Average Rating		
	Present	Absent	Total
Present	2	2	4
Absent	0	0	0
Total	2	2	4

Conclusion

The results confirm the prediction that when an average rating is present, people are much more focused during the search process and they do not search positive and negative reviews in a balanced way. The distribution presence leads to more search on the deviant reviews. People used both search aids as a basis for judgment of majority opinion and to form an initial impression. To a certain extent, the results also provide some support for the assumption that in the absence of search aids, people search reviews by date or randomly.

The results also reveal other considerations that people have during the search process. First, the distribution could serve as a basis for judgment of the average rating, when the average rating is absent. When the average rating is present, the presence of a distribution does not only divert people's attention away

from the average rating, it can also confirm the relevance of the average rating. This is due to the stimulus used in the current study (unimodal distribution with some dispersion), in which the mode of the distribution is largely consistent with the average rating. Second, negative reviews are more diagnostic and unexpected, thus it attracts more attention, especially when the average rating is absent. In the evaluation process, the average rating presence can shift people's attention away from counting the number of negative reviews only. This could be explained by the anchoring effect of the average rating presence. The results show that this anchoring bias can mitigate the negativity bias, when the average rating is high. That is, when the average rating is high and present, they search more consistent review and evaluate the resort based on the information gathered. These consistent reviews are positive.

Results of the Eye Movement Data Analysis

The results section comprises of three parts: first, preparation of the data for analysis and the usable sample are described. Second, the confirmatory factor analysis (CFA) results on the measurement model are presented. Lastly, the findings on the structural equation model are presented.

Data Preparation and Review Classification

Outliers were removed as follows. Cases with grand total search duration on the reviews (TSDR) greater than 300 seconds and grand total search duration on the Web site (TSDW) greater than 800 seconds were removed list-wise. Both cut-

off values were chosen based on the frequency histograms of the TSDR and TSDW. Responses with TSDR greater than 300 (3.08 standard deviations from the mean) and TSDW greater than 800 (3.12 standard deviations from the mean) were clearly separated from the rest of the responses on both histograms. After removing 6 outliers, the sample size of Study 2 was 312. Among the 312 responses, there are a total of 94 cases with systematic missing values (those participants that did not inspect the full view of at least one consistent and at least one deviant reviews). In SEM, responses with systematic missing values are removed list-wise and only 247 responses without any missing values are used for analysis. In sum, the sample in use comprises participants who view at least one consistent review and one deviant review in detail.

Measurement Model

A two-step approach was adopted in which first a valid and reliable measurement was established, and subsequently the structural model was tested using Mplus 6.12 (Muthén and Muthén 2011). To examine the measurement model, a confirmatory factor analysis (CFA) on the data collected from Study 2 was conducted. In particular, model fit was assessed in terms of four indices: comparative fit index (CFI), root mean square error of approximation (RMSEA), the Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). A model is considered to be satisfactory if CFI is greater than 0.95, TLI

is greater than 0.90, RMSEA is lower than 0.06, and SRMR is lower than 0.08 (Bearden, Netemeyer, and Mobley 1993; Hu and Bentler 1999).

The variables, the level of the average rating, presence of an average rating and presence of a distribution, are conceptually distinguishable from the self-reported measures and the measurement errors of these variables are zeros because these variables are manipulated based on the research design. Therefore, the reliability and discriminant validity of these variables do not constitute a concern. Thus, these variables were not included in the CFA model. Purchase intention was measured by a single indicator, “I will visit this resort for the coming holiday”, on a 7-point likert scale with anchors ranging from “very unlikely” to “very likely”. This item is similar to the Juster purchase probability scale (Juster 1966). For a latent variable with a single indicator, the square root of the reliability of the indicator can be used as the loading between the indicator and its latent variable (Bollen 1989; Kenny 1979; Ping 2004); and $(1 - \text{reliability})$ times the variance of the indicator can be used as the error variance of this single item measure (Bollen 1989). In this paper, the specifications of the single variable’s loading on the latent construct purchase intention and the error variance of this single item followed the above methods. According to Wright and MacRae (2007), the R-square of a linear regression of the predicted proportion on the actual proportion of purchase intention for Juster scale is .95. The value .95 was adopted as an estimate of reliability for purchase intention scale in the current study. All the other latent

variables were measured by multiple items. A number of control variables, including need for closure, need for cognition, knowledge about resorts, involvement in resorts, prior online review experience, and prior attitude towards the resort were also included in the CFA model. Particularly, need for closure consists of two distinct dimensions (Moneta and Yip 2004; Stalder 2009). The first dimension, need for structure, comprises three factors: discomfort with ambiguity, preference for order, and preference for predictability. Need for structure mainly relates to the freezing process—the inclination of individuals to accept the judgments they form rather than evaluate alternatives (Neuberg et al. 1997; Stalder 2007; Van Hiel, Pandelaere, and Duriez 2004). The most influential factor within this dimension is preference for predictability, which relates to one's dislike of disagreement within a group or unpredictable situations (Roets and Van Hiel 2011). The other two factors, discomfort with ambiguity and preference for order focus on the preference of a clear, expected and structured mode of life. Thus, they are comparatively less relevant to information search behavior and decision making. The second dimension relates to only one factor: decisiveness. Decisiveness mainly relates to the seizing process—the inclination of individuals to form rapid decisions or judgments. Thus, it is proposed to be one of the most important factors within need for closure that influences search behaviors. The two most influential factors (preference for predictability and decisiveness) within need

for closure were included in the model together with the other five control variables.

The results of CFA indicate that the initial measurement model did not fit the data well (χ^2 (696) = 1244.830, CFI = 0.867, TLI = 0.851, RMSEA = 0.057 and SRMR = 0.071). A careful inspection of the CFA output revealed that some items did not load highly on the designated latent factors (completely standardized loading lower than 0.50). To refine the measurement model, eleven items were dropped from prior attitude, attitude towards the resort, need for closure, involvement and prior online review experience (refer to the starred item in the appendix). The removed items were different from the retained items in terms of meaning and wording. For example, the fourth item of prior attitude—"I think it is good/bad" and the first item—"dislike/like" of product attitude were too general as compared to the rest of the items. The fourth item of prior attitude was reversely coded as compared to the rest of the items. The first item of product attitude, on the other hand, claims a much weaker stand than the other items, such as "disapprove/approve". Thus, both items were removed from the prior attitude scale and the post-search product attitude scale, respectively. The fifth item on preference for predictability scale—"I feel irritated when one person disagrees with everyone else in the group" was the only item based on group opinions. All the other four items on preference for predictability were based on general daily encounters instead of group scenarios. Thus, the fifth item on preference for

predictability scale was removed. In the decisiveness scale, the third and fourth items focused on the decision— “forming my own view” and post-decision feelings— “I feel relived”, while the first two items were about the eagerness of “finding a solution”. Therefore, only the first two items were retained in the decisiveness scale. In the involvement scale, the first to the third items focused on the image management aspect; and Items 4— “boring” and 8— “exciting” emphasized on the arousal aspect. These items differed from the from the rest of the items measuring involvement as "a person's perceived relevance” (Zaichkowsky 1985). Therefore, they were removed from subsequent analysis. Lastly, in the online review experience scale, the second item was on one’s history of searching online reviews, ranging from “never” to “been searching for a long time”. However, the first item was based on a general self-evaluated expertise and the last item was about the current frequency of usage. Both the first and the third items were very different from the second item in terms of meaning. Thus, the second item was removed from the online review experience scale.

After the measurement purification, a CFA was performed again with the remaining items. Compared to the previous model, the revised measurement model exhibits improved model fit (χ^2 (333) = 393.771, CFI = 0.982, the TLI = 0.978, RMSEA = 0.027, and SRMR = 0.042). The invariance of factor loadings, means, variances and covariances between the two groups (high level of average rating vs. low level of average rating) was also examined, by constraining the

corresponding parameters to be equal between the two groups. The invariance test is based on the difference in the χ^2 goodness-of-fit statistic between unconstrained and constrained CFA models. The test results show no significant differences in the factor loadings ($\Delta \chi^2 (19) = 7.458, p > .05$) and in the factor variances ($\Delta \chi^2 (10) = 14.762, p > .05$) between the two groups. However, the results do reveal significant differences in factor means ($\Delta \chi^2 (9) = 128.517, p < .05$) and covariances ($\Delta \chi^2 (45) = 82.508, p < .05$) between the factors. Therefore, a two-group CFA with factor loading and factor variance invariance (factor loadings and factor variances held equal between the two groups) was conducted. The CFA results are presented in Table 15. Particularly, the means of latent variables were set to zero in the first group (high average rating), and freely estimated in the subsequent group (low average rating). Because there is no unique solution for the means of the latent variables, the model estimates the differences between the means of the latent variables by group rather than values of the means of the latent variables for each group. In order to achieve model identification, the means for one of the groups are fixed to some arbitrary value, typically zero, to estimate the differences in means between groups.

Table 15

Estimated Factor Correlation Matrix Based on the Revised Measurement Model

High average rating		Mean difference ^b	SD	CR	AVE	1	2	3	4	5	6	7	8	9	10
1	PATT ^a	0.00	0.64	0.87	0.62	0.79 ^c									
2	ATT	0.00	1.06	0.86	0.61	0.23	0.78								
3	INT1	0.00	1.30	0.94	0.94	0.21	0.90	0.97							
4	CON	0.00	1.10	0.79	0.57	0.02	0.20	0.15	0.75						
5	DEC	0.00	1.01	0.79	0.65	0.05	0.04	0.00	-0.07	0.80					
6	PREDICT	0.00	0.90	0.71	0.38	0.03	0.00	-0.02	-0.16	0.50	0.62				
7	NC	0.00	0.98	0.76	0.52	0.01	-0.17	-0.16	0.02	0.17	0.65	0.72			
8	KNOWL	0.00	1.02	0.94	0.88	-0.06	0.13	0.12	-0.03	0.03	-0.07	-0.23	0.94		
9	INVOL	0.00	0.87	0.88	0.64	0.23	0.21	0.27	0.07	-0.11	-0.19	-0.25	0.37	0.80	
10	EXPERI	0.00	0.61	0.73	0.59	0.15	0.09	0.02	0.22	0.01	-0.15	-0.07	0.33	0.15	0.77
Low average rating		Mean difference ^b	SD	CR	AVE	1	2	3	4	5	6	7	8	9	10
1	PATT	0.05	0.64	0.83	0.56	0.75 ^c									
2	ATT	-1.81	1.06	0.84	0.58	0.12	0.76								
3	INT1	-2.28	1.30	0.97	0.97	0.13	0.90	0.98							
4	CON	0.21	1.10	0.80	0.57	-0.25	-0.49	-0.35	0.76						
5	DEC	-0.03	1.01	0.79	0.66	-0.02	-0.11	-0.07	0.14	0.81					
6	PREDICT	0.13	0.90	0.72	0.39	0.34	-0.03	-0.08	-0.04	0.48	0.62				
7	NC	0.33	0.98	0.76	0.52	0.28	0.14	0.13	-0.04	0.20	0.55	0.72			
8	KNOWL	-0.15	1.02	0.92	0.85	0.07	0.18	0.18	-0.03	-0.36	-0.15	0.00	0.92		
9	INVOL	-0.11	0.87	0.88	0.64	0.27	0.06	0.05	0.00	0.03	0.11	0.07	0.42	0.80	
10	EXPERI	-0.08	0.61	0.70	0.55	0.08	0.03	0.11	-0.05	0.01	-0.11	0.09	0.25	0.27	0.74

Notes. ^a PATT = prior attitude; ATT= product attitude (stage 2); INT1=purchase intention; CON=evaluation confidence; DEC=decisiveness; PREDICT= preference for predictability; NC= need for cognition; KNOWL=knowledge about resorts; INVOL=involved in resorts; EXPERI=online review experience; SD = standard deviations; CR=composite reliability; AVE=average variance extracted.

^b In order to achieve model identification, the means for one of the groups are fixed to zero, to estimate the differences in means between groups.

^c Value on the diagonal is the square root of AVE. Off-diagonal values represent correlations between the latent variables.

Results show that the differences between the means of post-search attitude and purchase intention of the two groups are -1.81 and -2.28 respectively. That is, in the high average rating group, participants' attitude and purchase intention are both higher than those in the low average rating group. This is consistent with our prediction. There are more positive reviews in the high average rating group. As one of the four components of search time, the amount of positive reviews available would contribute to the search time spent on positive reviews. Thus, the attitude and purchase intention in the high average rating group would become more positive. Additionally, preliminary analysis shows that the level of the average rating has a significant effect on attitude, on top of the search time's effects and regardless of the presence/absence of the average rating ($b = 1.623, p = .000$). People are exposed to the ratings accompanying each review. In the absence of the average rating statistics, they can still form their own judgement of the average rating, which in turn affects product attitude. Thus, the level of the average rating can affect post-search attitude.

In addition to the model fit, the reliability, the convergent validity, and the discriminant validity of the scale were examined. Table 15 describes the means, standard deviations, composite reliabilities (CR), and average variance extracted (AVE), and correlations of the factors based on the refined measurement model. Reliability was examined based on CR and AVE. A scale is said to be reliable if CR is greater than 0.70 and AVE is greater than 0.50 (Bagozzi and Yi 2012;

Bagozzi and Heatherton 1994). As shown in Table 15, the CRs range from 0.70 to 0.97. Most of the AVEs range from 0.52 to 0.97, except for the AVE of preference for predictability, which is 0.38 and 0.39 for the two groups respectively. On the other hand, convergent validity is established if all item loadings are equal to or above the recommended cut-off level of 0.50 (Hair et al. 1998). The CFA results show that all standardized item loadings are greater than 0.50, suggesting the convergent validity of the scale. Discriminant validity is the extent to which an item does not relate to the measures of other constructs. Discriminant validity is achieved if the square root of the AVE is larger than correlation coefficients (Fornell and Larcker 1981). All of the correlation estimates meet the criterion except in two cases. The exceptions are found between attitude and intention, and between preference for predictability (one dimension under need for closure) and need for cognition, when the average rating is high. This seems to pose less concern because the correlation between attitude and behavioral intentions has been well established in the literature (Petty and Krosnick 1995). In other words, both in the literature and in the current research, these two constructs are proposed to be strongly correlated. In addition, need for cognition has also been found to correlate with need for cognitive closure (Cacioppo et al. 1996). Therefore, it can be argued that in this study at least a reasonable extent of discriminant validity is established. Overall, the evidence of good model fit, reliability, convergent

validity, and discriminant validity indicates that the measurement model is appropriate for testing the structural model at a subsequent stage.

Structural Equation Model and Research Hypotheses

The two proposed Submodels A (search stage) and B (evaluation stage) were then tested separately through the structural equation modeling (SEM) technique using Mplus (Muthén and Muthén 2011). Submodel A depicts the influences of search aids on search variables and Submodel B focuses on how search time affects product attitude and purchase intention. The proposed SEM model consists of many variables and complex relationships. Testing the two submodels separately can help isolate any potential problems before testing the joint model. Based on the results of each submodel, a revised submodel was presented and tested respectively. As proposed in Hypothesis 5 under Submodel A, level of the average rating moderates the effects of average rating presence on the preview ratios. Besides, Hypothesis 8 and Hypothesis 9 under Submodel B propose a two-way interaction between search time and the level of average rating and a three-way interaction among search time, average rating presence and the level of average rating, respectively. One of the most common approaches of testing categorical moderator (i.e., interaction) hypotheses in SEM is by using multigroup SEM (Sauer and Dick 1993). Thus, two-group (a high average rating vs. a low average rating) SEM analysis was conducted both on the submodels and the joint models. Subsequently, the two revised submodels were combined into a revised joint

model—Model 1, representing the finalized results. Lastly, the revised joint model was compared with two competing models — Model 2 and Model 3, in order to validate the proposed model—Model 1.

Submodel A

Submodel A depicts the effects of search aids on four search variables, namely, the preview ratios of the consistent and deviant reviews, and the average search depth of the consistent and deviant reviews. The causal model of Submodel A was tested using the two-group SEM technique (a high average rating vs. a low average rating). As proposed in Hypothesis 5, the average rating level moderates the effects of average rating presence on the preview ratios. Thus, the coefficients of preview ratios on the average rating presence were unconstrained between the two groups (a high average rating vs. a low average rating). All the other regression coefficients were constrained to be equal between the two groups. Fit indices indicate that the model is a realistic representation of the data (χ^2 (726) = 852.058, CFI = 0.959, TLI = 0.952, RMSEA = 0.037, and SRMR = 0.060).

The initial Submodel A was compared with a constrained Submodel A, where all the regression coefficients were constrained to be equal between the two groups. The invariance test is based on the difference in the Chi square goodness-of-fit statistic between the partially unconstrained initial Submodel A and constrained Submodel A. No significant difference was found between the two models ($\Delta \chi^2$ (2) = 3.458, $p > .05$). That is, the moderation effects of the average

rating level on the effects of the average rating presence (Hypothesis 5) are not supported based on the Chi square test. Therefore, in the revised Submodel A, all the regression coefficients were constrained to be equal between the two groups.

The results of Submodel A also revealed some potential problems. For example, the results do not show the expected effects of the average rating presence, the distribution presence and the interaction between the two on the preview ratio of deviant reviews. The current research is mainly concerned about the effects of average rating presence, whereby people overly rely on one piece of information—the average rating—in the search process. Testing the relationships between the average rating presence and search behaviors, especially the search behaviors on consistent reviews, is the main focus of the study. Thus, the effect of the distribution presence and the average rating presence $\times$ distribution presence interaction effect on the preview ratio of deviant reviews were removed from the revised Submodel A, as both relationships were less essential to the research framework.

Most of the effects of control variables (i.e., decisiveness, preference for predictability, need for cognition, knowledge about resorts, involvement in resorts, prior online review experience, and prior attitude) on the four search variables across the two levels of the average rating are not significant. The only exception is that the effect of preference for predictability on preview ratio of deviant reviews is significant, when the average rating is low. To improve the results of the

Submodel A, all the insignificant control variables were removed in the revised Submodel A. Only preference for predictability was retained as a control variable in the revised Submodel A.

To summarize, in the revised Submodel A, all the regression coefficients were constrained to be equal between the two groups. The effect of the distribution presence and the average rating presence $\times$ distribution presence interaction effect on the preview ratio of deviant reviews were removed. All the control variables were also removed, except for preference for predictability.

The revised Submodel A was tested using the two-group SEM technique. Results indicate a better fit (χ^2 (118) = 135.111, CFI = .985, TLI = 0.981, RMSEA = 0.034, and SRMR = 0.060). All of the relationships in the revised Submodel A remain largely the same, except for some minor changes in the path coefficients or p-values. No significant effect was found on the control variable, preference for predictability. Thus, it was subsequently removed in the joint model.

Submodel B

The causal model of Submodel B was also tested using the two-group (a high average rating vs. a low average rating) SEM technique. The proposed Submodel B involves the interaction between attitude and evaluation confidence, which are measured by multiple items. A parsimonious estimation technique using the product of summed indicants by Ping (1995) was used to estimate the interaction effect involving latent variables measured by multiple items (the attitude $\times$

evaluation confidence interaction effect on purchase intention). In addition, Hypotheses 8 and 9 propose a two-way interaction between search time and the level of the average rating, and a three-way interaction among search time, average rating presence and the level of the average rating, respectively. Both interaction effects were tested using two-group SEM (a high average rating vs. a low average rating), where the coefficients of the effects of search time, and of the search time $\times$ average rating presence interaction effects on attitude were unconstrained between the two groups. All the other regression coefficients were constrained to be equal between the two groups. The results of SEM indicate that the initial Submodel B does not fit the data well (χ^2 (263) = 810.658, CFI = 0.723, TLI = 0.698, RMSEA = 0.130, and SRMR = 0.140).

The results show that the average rating presence did not enhance the effects of search time spent on consistent reviews on product attitude, regardless of the level of the average rating. To improve the overall model fit, the search time $\times$ average rating presence interaction effects were removed from the revised Submodel B.

Results do not support the positive effects of both average rating presence and the distribution presence on the evaluation confidence. The evaluation confidence also does not strengthen the positive effect of attitude. To reinforce the main focus of the study and to improve the overall model fit of Submodel B, the factor confidence and the non-significant confidence $\times$ attitude interaction effect

were removed from the revised Submodel B. The effects of search aids (average rating and distribution presence) on confidence were also removed.

Interestingly, results of the initial Submodel B show some similarities in the coefficient of time spent on consistent reviews on product attitude in the high average rating group and the coefficient of time spent on deviant reviews on attitude in the low average rating (positive reviews). Similarities are also found between the coefficient of time spent on deviant reviews on attitude in the high average rating group and the coefficient of time spent on of consistent reviews on attitude in the low average rating (negative reviews). Thus, a chi square difference test is conducted to compare these two pairs of coefficients, respectively. Specifically, the initial Submodel B was compared with another version of Submodel B, with the following constraints applied:(a) the coefficient of time spent on consistent reviews on product attitude in the high average rating group and the coefficient of time spent on deviant reviews on product attitude in the low average rating (positive reviews) are constrained to be equal; and (b) the coefficient of time spent on deviant reviews on product attitude in the high average rating group and the coefficient of time spent on of consistent reviews on product attitude in the low average rating (negative reviews) are constrained to be equal. The test results show no difference between the two submodels ($\Delta \chi^2 (2) = 0.699, p > .05$), indicating the above two coefficients within each sets of are equal. Based on the broad categorization of the reviews, the consistent reviews in the high average rating

group and the deviant reviews in the low average rating group represent the same set of reviews—positive reviews with rating from 6 to 10. Similarly, the deviant reviews in the high average rating group and the consistent reviews in the low average rating group are the same set of reviews—negative reviews with rating from 1 to 5. Thus, the results show that the effects of search time on attitude are related to the valence of the reviews rather than the consistency of the reviews with the average rating.. Therefore, the constrained relationships are adopted in the revised Submodel B and revised joint model.

The control variable prior attitude had a positive effect on product attitude, when the average rating was high ($b=0.462$, $p=.005$). Thus, prior attitude is retained as a control variable in the revised Submodel B.

To summarize, in the revised Submodel B, the search time $\times$ average rating presence interaction effect was removed. In addition, the factor evaluation confidence and the confidence $\times$ attitude interaction effect were also removed from the revised Submodel B. The effects of search aids (average rating and distribution presence) on confidence were also removed. The coefficient of time spent on consistent reviews on product attitude in the high average rating group and the coefficient of time spent on deviant reviews on product attitude in the low average rating group (positive reviews) are constrained to be equal and vice versa (negative reviews).

The revised Submodel B was tested using the two-group (a high average rating vs. a low average rating) SEM technique. Fit indices indicate that the model is a much better representation of the data as compared to the initial Submodel B (χ^2 (99) =114.487, CFI = 0.987, TLI = 0.986, RMSEA = 0.036, and SRMR=0.062). All of the relationships in the revised Submodel B remained largely the same, except for some minor changes in the path coefficients or p-values.

Joint Model

Results on the initial and revised joint models are presented. The initial joint model represents the original hypotheses and a combination of the initial Submodels A and B. The revised joint model represents a combination of the revised Submodels A and B.

The initial joint model was tested using the two-group (a high average rating vs. a low average rating) SEM technique. This model represents the original hypotheses and is basically a combination of the initial Submodels A and B, without the control variables in the initial Submodel A. Because the Submodel A analysis shows that the effects of the control variables on the search variables are nonsignificant, these effects are excluded from the joint model. The Submodel B analysis shows that the effect of prior attitude on the post search product attitude is significant. Thus, prior attitude was retained as a control variable on the product attitude. Fit indices indicate that the initial joint model did not represent the data

well (χ^2 (447) = 1170.463, CFI = 0.765, TLI = 0.740, RMSEA = 0.114, and SRMR=0.122). The results are largely consistent with the initial findings from Submodels A and B. The path coefficients of the initial joint model are presented in Table 16.

The revised Submodels A and B were combined into the revised joint model. The revised joint model was tested using the two-group (a high average rating vs. a low average rating) SEM technique. Fit indices indicate that the model is a realistic representation of the data (χ^2 (269) = 313.937, CFI = .979, TLI = 0.977, RMSEA = 0.037, and SRMR=0.069). All of the significant relationships remain largely unchanged. The path coefficients of the initial and revised joint models are presented in Table 16. Figure 10 reports the results of the revised joint model. The hypothesis test results are summarized in Table 17.

Figure 10
Unstandardized Coefficient Estimates of the Revised Joint Model

Notes. ^a When two coefficient estimates are presented for both high and low average rating groups, the numbers in the brackets pertain to the low average rating group.

^b Bold coefficients indicate support for hypothesis.

^c !p<.1, *p<.05, **p<.01, ***p<.001 (two-tailed).

Table 16
Unstandardized Coefficient Estimates of the Initial and Revised Joint Models

<i>Search Process</i>	Initial Joint Model		Revised Joint Model	
	Low average rating	High average rating	Low average rating	High average rating
Preview ratio of consistent reviews^a	R ² = .022 R ² _{adj.} = .010	R ² = .028 R ² _{adj.} = .016	R ² = .023 R ² _{adj.} = .011	R ² = .029 R ² _{adj.} = .017
Average rating presence	0.170!	0.170!	0.158! ^{b,c}	0.158!
Distribution presence	0.088	0.088	0.050	0.050
Average rating presence × Distribution presence	-0.314*	-0.314*	-0.291*	-0.291*
Preview ratio of deviant reviews	R ² = .004 R ² _{adj.} = -.008	R ² = .007 R ² _{adj.} = -.005	R ² = .002 R ² _{adj.} = -.002	R ² = .004 R ² _{adj.} = .000
Average rating presence	-0.038	-0.038	-0.071	-0.071
Distribution presence	0.099	0.099	-- ^d	--
Average rating presence × Distribution presence	-0.061	-0.061	--	--
Average search depth of consistent reviews	R ² = .035 R ² _{adj.} = .027	R ² = .033 R ² _{adj.} = .025	R ² = .035 R ² _{adj.} = .027	R ² = .033 R ² _{adj.} = .025
Average rating presence	0.24*	0.24*	0.24*	0.24*
Distribution presence	-0.151	-0.151	-0.151	-0.151
Average search depth of deviant reviews	R ² = .020 R ² _{adj.} = .012	R ² = .032 R ² _{adj.} = .024	R ² = .020 R ² _{adj.} = .012	R ² = .032 R ² _{adj.} = .024
Average rating presence	0.257*	0.257*	0.257*	0.257*
Distribution presence	0.094	0.094	0.094	0.094

	Initial Joint Model		Revised Joint Model	
<i>Search Process (Cont')</i>	Low average rating	High average rating	Low average rating	High average rating
Search time on consistent reviews	R ² = .593 R ² _{adj.} =.590	R ² = .676 R ² _{adj.} =.673	R ² = .593 R ² _{adj.} =.590	R ² = .676 R ² _{adj.} =.673
Preview ratio of consistent reviews	0.577***	0.621***	0.577***	0.621***
Average search depth of consistent reviews	0.968***	1.093***	0.968***	1.093***
Search time on deviant reviews	R ² = .868 R ² _{adj.} =.867	R ² = .833 R ² _{adj.} =.832	R ² = .868 R ² _{adj.} =.867	R ² = .832 R ² _{adj.} =.832
Preview ratio of deviant reviews	0.799***	0.726***	0.799***	0.726***
Average search depth of deviant reviews	0.992***	0.967***	0.992***	0.967***
<i>Evaluation Process</i>				
Product attitude	R ² = .156 R ² _{adj.} =.138	R ² = .107 R ² _{adj.} =.088	R ² = .116 R ² _{adj.} =.100	R ² = .118 R ² _{adj.} =.104
Search time on consistent reviews	-0.326***	0.277*	-0.326***	0.173**
Search time on deviant reviews	0.173***	-0.325**	0.173***	-0.326***
Search time on consistent reviews× Average rating presence	-0.067	-0.183	--	--
Evaluation confidence	R ² = .015 R ² _{adj.} =.007	R ² = .016 R ² _{adj.} =.008		
Average rating presence	-0.184	-0.184 ^c	--	--
Distribution presence	0.203	0.203	--	--
Purchase intention	R ² = .838 R ² _{adj.} =.836	R ² = .856 R ² _{adj.} =.854	R ² = .796 R ² _{adj.} =.795	R ² = .823 R ² _{adj.} =.822
Product attitude	1.188***	1.188***	1.107***	1.107***
Evaluation confidence	0.115	0.115	--	--
Product attitude× Evaluation confidence	-0.011	-0.011	--	--

Note. ^a The variable in bold is regressed on the following variables.

^b Bold coefficients indicate support for hypothesis.

^c !p<.1, *p<.05, **p<.01, *** p<.001 (two-tailed).

^d The relationship is not included in the revised joint model.

Table 17
Hypotheses Testing

	Hypothesis	Result
Search Process		
H1(a)	Provision of an average rating increases the preview ratio of consistent reviews, when the distribution is absent.	Supported
H1(b)	Provision of an average rating decreases the preview ratio of deviant reviews, when the distribution is absent.	Not Supported
H2	Provision of a product rating distribution decreases (a) the positive impact of an average rating provision (i.e., making the impact less positive) on the preview ratio of consistent reviews, and (b) the negative impact of an average rating provision (i.e., making the impact less negative) on the preview ratio of deviant reviews.	Partially supported
H3	Provision of a product rating distribution does not affect preview ratios of (a) the consistent reviews or (b) the deviant reviews, when the average rating is absent.	Supported
H4(a)	Provision of a product rating distribution decreases the preview ratio of consistent reviews, when the average rating is present.	Supported
H4(b)	Provision of a product rating distribution increases the preview ratio of deviant reviews, when the average rating is present.	Not Supported
H5	A low level of average rating magnifies (a) the positive effect of provision of an average rating on the preview ratio of the consistent reviews, and (b) the negative effect of provision of an average rating on the preview ratio of the deviant reviews.	Not Supported
H6	Provision of an average rating decreases the average search depth of both (a) consistent and (b) deviant reviews.	Not Supported (significant but in the opposite direction)
H7	Provision of a product rating distribution increases the average search depth of both (a) consistent and (b) deviant reviews.	Not Supported
Evaluation Process		
H8a	Search time on consistent reviews increases product attitude when the average rating is high, and decreases product attitude when the average rating is low.	Supported
H8b	Search time on deviant reviews decreases product attitude when the average rating is high, and increases product attitude when the average rating is low.	Supported
H9	Provision of an average rating increases (magnifies) the positive effect of search time on consistent reviews on product attitude when the average rating is high.	Not Supported
H10	Provision of search aids increases evaluation confidence as compared to the conditions when the search aids are not provided.	Not Supported
H11	The evaluation confidence strengthens the positive relationship between attitude and purchase intention.	Not Supported

SEM Results of Alternative Specifications of Consistent and Deviant Reviews

SEM results based on two alternative specifications of consistent and deviant reviews are also provided in Appendix D. The original results (current scheme) were labeled as Classification One. Based on Classification One, when the average rating is high, reviews with rating 6 to 10 are classified as the consistent reviews and others are classified as the deviant reviews. When the average rating is low, reviews with rating 1 to 5 are classified as the consistent reviews and others are classified as the deviant reviews. Classifications Two and Three adopted some alternative methods of specifications of consistent and deviant reviews. Based on Classification Two, when the average rating is high, reviews with rating 7 to 8 are classified as the consistent reviews and others are classified as the deviant reviews. When the average rating is low, reviews with rating 3 to 4 are classified as the consistent reviews and others are classified as the deviant reviews. Based on Classification Three, when the average rating is high, reviews with rating 6 to 9 are classified as the consistent reviews and others are classified as the deviant reviews. When the average rating is low, reviews with rating 2 to 5 are classified as the consistent reviews and others are classified as the deviant reviews.

Fit indices of both Classification Two (χ^2 (269) = 367.097, CFI = 0.944, TLI = 0.937, RMSEA = 0.059, and SRMR = 0.086) and Three (χ^2 (269) =

336.891, CFI = 0.969, TLI = 0.965, RMSEA = 0.045, and SRMR= 0.077) indicate that the original model is a better representation of the data.

Under Classification Two scheme (narrowest scheme), none of the proposed effects of search aids were significant ($p > .1$) in the search stage. In the evaluation stage, only the consistent reviews under low average rating condition and the deviant reviews under high average rating condition significantly affect attitude ($b = -0.201$, $p = .005$). The results provide some partial support for H8. None of the other hypotheses are supported using Classification Two.

When a relatively broader classifying scheme was adopted under Classification Three (middle scheme), the results greatly improved. In the search stage, the negative interaction effect between average rating presence and distribution presence on preview ratio of consistent reviews is marginally significant ($b = -0.162$, $p = .092$), supporting H2a. In the evaluation stage, results show that search time spent on consistent reviews has a positive effect on product attitude when the average rating is high ($b = 0.129$, $p = .035$), and a negative effect on product attitude when the average rating is low ($b = -0.273$, $p = .000$). Search time spent on deviant reviews has a negative effect on product attitude when the average rating is high ($b = -0.273$, $p = .000$), and a significant positive effect on product attitude when the average rating is low ($b = 0.129$, $p = .035$). Hypotheses 8a and 8b are therefore supported under Classification Three scheme.

The broadest classification under Classification One showed the best results among the three classifications. Many of the major hypotheses are supported as presented in the results section. In sum, based on both the parameter estimates and the fit index, Classification One is a better representation of the data. Thus, Classification One is chosen eventually.

Effects of Average Rating Presence on Preview Ratios

Hypothesis 1a, 2a, 3 and 4a describes average rating presence $\times$ distribution presence interaction effect on the preview ratio of consistent reviews. Hypothesis 1a proposes that provision of an average rating increases the preview ratio of consistent reviews, when the distribution is absent. The effects of average rating presence are examined under the presence and absence of the distribution respectively, as a significant interaction was found between the two factors. The p-values of the simple main effects were generated by changing the reference group of distribution presence through recoding the variable. Specifically, to estimate the effects of average rating presence under the presence of distribution, the distribution presence variable was coded with the distribution presence as the reference group, that is, zero for presence and one for absence. Subsequently, to estimate the effects under the absence of distribution, the distribution presence variable was coded with the distribution absence as the reference group, that is, zero for absence and one for presence. The analysis of the simple main effects reveals that when the distribution is absent, the average rating presence has a

marginally significant positive effect on the preview ratio of consistent reviews ($b = 0.158, p = .085$). When the distribution is present, the average rating presence does not affect the preview ratio of consistent reviews ($b = -0.133, p = .132$). Thus, Hypothesis 1a is supported.

Provision of a product rating distribution was proposed to decrease the positive effect of average rating presence on the preview ratio of consistent reviews in Hypothesis 2a. The results pertaining to the interaction support Hypothesis 2a ($b = -0.291, p = .014$). Figure 11 shows the mean preview ratios of the consistent reviews under the eight different conditions. When the average rating is present and the distribution is present, the mean preview ratio of consistent reviews is the highest. When both search aids are present, the mean preview ratio of consistent reviews is the lowest. The findings are consistent with the interaction results. When the distribution is absent, the mean under the average rating presence condition is higher than the mean under the average rating absence condition. When the distribution is present, the mean under the average rating presence condition is lower than the mean under the average rating absence condition.

Figure 11
Marginal Means of Log-transformed Preview Ratio of the Consistent Reviews

(a) *Low average rating*

(b) *High average rating*

Hypothesis 1b proposes that provision of an average rating decreases the preview ratio of deviant reviews, when the distribution is absent. Results do not support Hypothesis 1b ($b = -0.071, p = .406$). In addition, Hypothesis 2b describes the impact of average rating presence on the preview ratio of deviant reviews when the distribution is present. No significant effect was found on the average rating presence $\times$ distribution presence on the preview ratio of deviant reviews. This interaction term was removed from the revised joint model (see results in initial joint model in Table 16, $b = -0.061, p = .719$).

The Effects of Average Rating Presence on Average Search Depth

Hypotheses 6a and 6b describe the negative effect of average rating presence on the average search depth of consistent and deviant reviews. However, results show that the average rating presence had a significant positive effect on the average search depth of both consistent reviews ($b = 0.240, p = .013$) and deviant reviews ($b = 0.257, p = .014$). Hypotheses 6a and 6b are not supported. These two relationships are in the opposite directions as compared to the proposed model.

Effects of the Distribution Presence on Preview Ratios and Average Search Depth

Hypothesis 3a posits that provision of a product rating distribution does not affect preview ratios of the consistent reviews, when the average rating is absent. On the other hand, hypothesis 4a proposes that provision of a product rating distribution decreases the preview ratio of consistent reviews, when the average

rating is present. The effects of distribution presence are examined under the presence and absence of the average rating respectively, as a significant interaction was found between the two factors. The p-values of the simple main effects were generated by changing the reference group based on the average rating presence through recoding the variable. Precisely, to estimate the effects of the distribution presence under the presence of average rating, the average rating presence variable was coded with the average rating presence group as the reference group, that is, zero for presence and one for absence. Subsequently, to estimate the effects under the absence of average rating, the average rating presence variable was coded with the average rating absence as the reference group, that is, zero for absence and one for presence. An analysis of the simple main effects reveals that when the average rating is present, the distribution presence has a negative effect on the preview ratio of consistent reviews ($b = -0.241, p = .004$). When the average rating is absent, the distribution presence has no effect on the preview ratio of consistent reviews ($b = 0.050, p = .549$). Therefore, Hypothesis 3a and 4a are supported.

Hypothesis 3b posits that provision of a product rating distribution does not affect preview ratios of the deviant reviews, when the average rating is absent. Hypothesis 4b describes the positive effect of the product rating distribution presence on the preview ratio of deviant reviews, when the average rating is present. Results show that the distribution presence do not affect the preview ratio of deviant reviews ($b = 0.099, p = .408$). Thus, Hypothesis 4b is not

supported. In addition, the rating distribution presence $\times$ average rating presence interaction effect on the preview ratio of deviant reviews is also not significant ($b=-0.061, p = .719$). The results provide some support for Hypothesis 3b. These two relationships were removed from the revised joint model (see results of the initial joint model in Table 16). The nonsignificant results could be due to the test stimulus used. A unimodal distribution with some dispersion was used as the test stimulus in the study to represent a “dispersed distribution” due to its pervasion. In statistics, dispersion is the variability or spread in a variable. Applied to the current context, the dispersion reflects the heterogeneity of opinions. The level of dispersion of the test stimulus used might not be sufficient to drive the effects. That is, owing to the insufficient level of dispersion, the unimodal distribution does not effectively highlight the discrepant ratings. Therefore, the distribution presence does not direct people’ attention to these deviant reviews.

Hypotheses 7a and 7b describes the positive effects of product rating distribution presence on the average search depth of both consistent and deviant reviews. The results show that the distribution presence has no effect on the average search depth of consistent reviews ($b=-0.151, p = .115$) or deviant reviews ($b=0.094, p = .366$). Therefore, hypotheses 4a and 4b are not supported.

Effects of the Preview Ratios and Average Search Depth on Search Time

The results show significant positive effects of the preview ratios ($b= 0.621, p = .000$) and average search depth of consistent reviews ($b= 1.093, p = .000$) on

the search time spent on the consistent reviews. Results also show significant positive effects of the preview ratios ($b = 0.577, p = .000$) and average search depth of deviant reviews ($b = 0.968, p = .000$) on the search time spent on deviant reviews. The results are consistent with the decomposition of search time.

Search Time $\times$ Level of the Average Rating Interaction Effect on Product Attitude

Hypotheses 8a and 8b describes the effects of search time on consistent or deviant reviews on product attitude. Results show that search time spent on consistent reviews has a positive effect on product attitude when the average rating is high ($b = 0.173, p = .006$), and a negative effect on product attitude when the average rating is low ($b = -0.326, p = .000$). Search time spent on deviant reviews has a negative effect on product attitude when the average rating is high ($b = -0.326, p = .000$), and a significant positive effect on product attitude when the average rating is low ($b = 0.173, p = .006$). Hypotheses 8a and 8b are therefore supported. Interestingly, the parameter estimate of the impact of search time on consistent reviews on attitude under high average rating condition equals to the parameter estimate of the impact of search time on deviant reviews on attitude under low average rating condition and vice versa ($\Delta \chi^2 (2) = 0.187, p > .05$). This reveals the fact that the valence of the reviews is the main driving factor of attitude instead of the classification between the consistent and deviant reviews.

Hypothesis 9 proposes that provision of average rating increases the effect of search time allocated to consistent reviews on attitude. The results show that the average rating presence does not enhance the effects of search time spent on consistent reviews on product attitude, when the average rating is high ($b = -0.183$, $p = .304$) or when the average rating is low ($b = -0.067$, $p = .703$). Hypothesis 9 is not supported. This relationship was subsequently removed from the revised joint model (see results in initial joint model in Table 16). One possible explanation for the lack of support is that when a consumer reads the reviews, she is influenced by both the review ratings and the content. She can still form a temporary judgment of the average rating, even if the average rating is absent. This temporary judgment is likely to have the same valence as the average rating based on majority opinion. She could still organize the product-related information in memory around the temporary judgment and draw on that information in a later decision. Therefore, this study does not find any evidence for the interaction effect between the search time spent on consistent reviews and the average rating presence on product attitude.

Product Attitude on Purchase Intention

Product attitude has a positive effect on the purchase intention as proposed ($b = 1.107$, $p = .000$). Hypothesis 10 describes the positive effects of provision of search aids on evaluation confidence and Hypothesis 11 describes the evaluation confidence $\times$ attitude interaction. However, results did not provide support for the

positive effect of both average rating presence ($b = -0.184$, $p = .233$) and the distribution presence ($b = 0.203$, $p = .197$) on the evaluation confidence. Therefore, Hypothesis 10 is not supported. Both relationships were removed from the revised joint model (see results in initial joint model in Table 16). The evaluation confidence also did not strengthen the positive effect of attitude ($b = -0.010$, $p = .330$). As such, Hypothesis 11 is not supported. This relationship was also removed from revised joint model (see results in initial joint model in Table 16). Although the moderating role of confidence on attitude–intention relationship is well established in the literature (Ajzen 1988; Kraus 1995; Petkova et al. 1995), some other researchers do hold a different view. According to Howard and co-workers (Howard 1989; Howard, Shay, and Green 1988; Howard and Sheth 1969), confidence also has a causal influence on purchase intention – that is, confidence is a predictor of purchase intention. A number of studies support this relationship (Howard et al. 1988; Laroche, Kim, and Zhou 1996; Laroche and Sadokierski 1994). One possible explanation for these conflicting views lies in differences in the methodology of the two types of study. The ‘predictor’ studies were based on large-sample consumer surveys, where confidence was measured (Howard et al. 1988; Laroche et al. 1996; Laroche and Sadokierski 1994), while the ‘moderator’ studies were based on laboratory experiments, where confidence was manipulated (Berger and Mitchell 1989; Tormala, Clarkson, and Petty 2006). The results in the initial joint model were consistent to the “predictor” view of confidence to some

extent as the statistical significance of the direct effect of confidence on intention was close to .10 ($b=0.115, p=.176$).

The control variable, prior attitude towards the product, has a positive effect on product attitude only when the average rating is high ($b=0.451, p=.005$), but not when the average rating is low ($b=0.149, p=.305$).

Model Validation

To further validate the revised joint model (Model 1), Model 1 was compared with a competing model – Model 2. Model 1 proposes that search variables fully mediate the effects of search aids on product attitude. The competing model – Model 2 is an extended model, which includes all the relationship in Model 1 and additional direct effects from search aid variables (average rating presence, distribution presence and average rating presence $\times$ distribution presence) on product attitude. The three additional free coefficients were constrained to be equal between the two groups (high vs. low average rating). Model 1 is nested within Model 2. A Chi square difference test is used to compare the goodness of fit between Models 1 and 2. No significant difference was found between the extended model (Model 2) and the revised joint model (Model 1) ($\Delta \chi^2 (3) = 7.456, p > .05$). This suggests that the proposed search variables' full mediation of search aids' effects on product attitude is a good representation of the data, although the direct effect of certain search aids (e.g., the rating distribution)

on the attitude is also found marginally significant within the extended model ($b=0.321, p=.081$).

In addition, the three indirect effects, namely, e rating distribution) on the attitude is also found marginally significant with $\text{sence} \rightarrow \text{search variables} \rightarrow \text{attitude}$, “and the average rating presence $\times$ distribution presence $\rightarrow$ search variables $\rightarrow$ attitude” are also estimated under different conditions (average rating present vs. absent; distribution presence vs. absent) using two different methods to check the robustness of the results across methods: a two-group SEM (high level of average rating vs. low level of average rating) and a two-group bootstrapping analysis using Mplus. The estimated indirect path coefficients using two-group SEM are shown in Table 18. According to Zhao, Lynch and Chen (2010), bootstrapping method is relatively high in power in the test for mediation as compared to other methods (e.g., Sobel test). The 95% confidence intervals of the parameter estimates using bootstrapping method are shown in Appendix E (Table 20). The results based on both methods are largely consistent with each other, except for two cases; (a) the total indirect effects of distribution presence to attitude under the condition that the average rating is absent, when the average rating is high; and (b) the total indirect effects of distribution presence to attitude under the condition that the average rating is absent, when the average rating is low. According to the two-group bootstrapping results, the coefficient estimates of the total indirect effects of distribution presence to attitude under the condition that the

average rating is absent are insignificant at 95% confidence level (two-tailed). However, based on the two-group SEM results, both coefficient estimates were marginally significant. In sum, the two-group bootstrapping results largely confirmed the two-group SEM results. The next section contains a detailed discussion on each of the indirect effect based on the two-group SEM results.

First, refer to the results on the indirect effects of average rating. Results show that when the average rating is low, presence of an average rating has a marginally significant negative total effect on product attitude ($b = -0.071$, $p = .063$), under the condition that the distribution is absent. Specifically, this negative effect is contributed mainly by the “average rating presence → average search depth of consistent reviews → search time on consistent reviews → attitude” relationship ($b = -0.076$, $p = .033$) and the presence → average search depth of consistent reviews → search time on consistent reviews → attitude” relationship (effect of an ($b = -0.030$, $p = .122$)). The effects by the average search depth of consistent and deviant reviews are in the opposite direction. The effects cancel out under the high average rating condition. Under the low average rating condition, the effects do not cancel out completely.

Secondly, referring to the indirect effects of distribution presence, when the average rating is low, distribution presence has a positive total effect on attitude, when both the average rating is absent ($b = 0.054$, $p = .095$) and present ($b = 0.109$, $p = .006$). When the average rating is high, distribution presence has a negative

total effect on attitude when both the average rating is absent ($b = -0.053, p = .096$) and present ($b = -0.084, p = .026$). When the average rating is absent, the effects are mainly contributed by the average search depth of consistent reviews. When the average rating is present, the effects are mainly contributed by the preview ratios of consistent reviews.

In sum, across the two stages (search and evaluation), the total indirect effects of average rating presence depend on the level of the average rating and the absence of a rating distribution. On the other hand, the presence of a distribution shifts the attitude in a direction opposite to the valence of the average rating.

Table 18
Unstandardized Estimates of the Indirect Effects of Average Rating Presence and Distribution Presence

	Low average rating		High average rating	
	Distribution absent	Distribution present	Distribution absent	Distribution present
Effects from Average Rating Presence to Attitude				
<i>Total indirect effects</i>	-0.071! ^a	-0.016	-0.002	-0.033
<i>Specific indirect effects</i>				
average rating presence→ preview ratio of consistent reviews→ search time on consistent reviews→ attitude	-0.030	0.025	0.017	-0.014
average rating presence→ average search depth of consistent reviews→ search time on consistent reviews→ attitude	-0.076* ^b		0.045!	
average rating presence→ preview ratio of deviant reviews→ search time on deviant reviews→ attitude	-0.010		0.017	
average rating presence→ average search depth of deviant reviews→ search time on deviant reviews→ attitude	0.044!		-0.081*	
	Low average rating		High average rating	
	Average Rating absent	Average Rating present	Average Rating absent	Average Rating present
Effects from Distribution Presence to Attitude				
<i>Total indirect effects</i>	0.054!	0.109*	-0.053!	-0.084*
<i>Specific indirect effects</i>				
distribution presence→ preview ratio of consistent reviews→ search time on consistent reviews→ attitude	-0.009	0.045*	0.005	-0.026!
distribution presence→ average search depth of consistent reviews→ search time on consistent reviews→ attitude	0.048		-0.029	
distribution presence→ average search depth of deviant reviews→ search time on deviant reviews→ attitude	0.016		-0.030	

Note. ^a!p<.1, *p<.05 (two-tailed).

^b Only one coefficient estimate is presented when the effects are the same across the two conditions.

CHAPTER 6

DISCUSSION

This paper provides insights into both the search and online review literature through its conceptualization, methodology and findings. First, this study breaks down the search time to the selection and processing components and classifies the reviews as consistent and deviant reviews. Second, the study expands the research scope to a large amount of reviews under each discussion and incorporates process-tracing measures in online review research. Lastly, the findings provide insights into the complex relationships between search aids, search variables, product attitude and purchase intention during both search and evaluation stages. Each of the contributions is discussed in the next section.

In contrast to previous information search empirical studies in which search time was used as a key variable of interest (Ratchford and Srinivasan 1993), this study expands the traditional measures of search time to the online review search context with large amounts of reviews and breaks it down to selection and processing components. Results show that search aids exert different effects on the two distinct search components: preview ratio (selection component) and average search depth (processing component). These findings thus validate the usefulness of such decomposition.

In addition, although the classification of affirming and dissenting messages has been developed in research on social groups (Matz and Wood 2005),

it has not been applied to the online review context. Thus, this paper also makes a contribution by refining and extending the above distinction under the online review research context and proposes the classification of consistent and deviant reviews. Furthermore, search aids are found to exert different effects on these two distinct sets of review messages. Such findings validate the usefulness of the categorization.

Furthermore, prior individual-level research on online reviews mainly focuses on the information source (e.g., the reviewers), the message characteristics, and the information receivers' influence of consumer purchase decisions based on one or a few messages (Petty and Cacioppo 1996; Zhang and Watts 2008). Previous studies on the online reviews market-level analysis focus mainly on examining the total effects of the average ratings, volumes and the valence of consumer reviews on product sales (Chevalier and Mayzlin 2006; Godes and Mayzlin 2004; Liu 2006; Trusov et al. 2009). The current study expands the research scope of individual-level analysis to a large amount of review information under each discussion, and examines the impact of the forum search aids and the mediating roles of search variables in such context.

Also, most prior individual-level WOM or online review research are based on self-reported attitude/intention data (Cheung et al. 2009; Petty and Cacioppo 1996; Zhang and Watts 2008). These studies do not incorporate process-tracing measures (e.g., eye movement measures) in their examination. Thus, they

do not capture the underlying process regarding how consumers search these reviews. As a result, the processes through which the rating statistics affect consumers' buying behaviour are little known and not empirically confirmed. Therefore, this study helps fill up this knowledge gap. In the current research, the search variables are empirically captured and analyzed using consumers' eye movements over a threaded discussion. The study measures the consumers' eye movements while processing the reviews, including the exact area that the consumer is paying attention to and the amount of attention paid while processing the reviews. In order to test the proposed model, this process tracing method is useful as it enables accurate measurements of the number of reviews previewed, the number of reviews selected to read and the time spent reading them within each set of reviews. Besides, qualitative measures, including concurrent and retrospective think-aloud, are used to complement the quantitative measures. Thus, the findings can add new knowledge to online review research.

The findings provide insights into the complex relationships between search aids, search variables, product attitude and purchase intention during both search and evaluation stages. In the search stage, the results show that presence of an average rating has a positive effect on the two major components of search time, namely, the preview ratio and average search depth of the consistent reviews. That is, when an average rating is present on the online review website, consumers preview more consistent reviews and spend more time reading each of them in

detail. Thus, their product attitude and purchase intention become more in line with the provided average rating, when the average rating is provided. These findings provide support to the assimilation effect in anchoring literature (Epley and Gilovich 2006; Tversky and Kahneman 1992), whereby the average product rating presented to consumers serves as an “anchor”, attracting attention to the consistent reviews. Presence of a product rating distribution can mitigate such positive effect on the preview ratio of the consistent reviews, supporting the forewarning effect (Highhouse and Bottrill 1995). While anchoring effect has been previously examined in decision making context, the current study shows these findings in both information search stage and decision making stage. Interestingly, presence of an average rating also has a positive effect on the average search depth of deviant reviews. Such findings can be explained by the other alternative effects proposed earlier based on the selective attention theory in information processing literature (Treisman 1979). A deviant review becomes especially distinctive, when the average rating is present. Thus, consumers will examine this discrepant position carefully in order to understand why some would be willing to deviate (Moskowitz 1996).

This study also finds some support for the effects of presence of a distribution on search behaviors. The simple main effect of distribution on preview ratio of consistent reviews is negative, on condition that the average rating is provided. Such effect suggests that distribution reduces selection of consistent

reviews when average rating is present. This is consistent with the prediction that a distribution directs people's attention away from the reviews that are consistent with the average rating, if the average rating is present. When the average rating is absent and the distribution is present, people can still form their own judgment of the average rating based on the distribution. Therefore, when the mode of the distribution is consistent with the average rating, the presence of the distribution can direct some attention to the consistent reviews and cancel out its negative effect on these reviews. Thus, the negative effect of distribution presence is not found under the condition that the average rating is absent.

However, results do not support the notion that presence of a distribution with dispersion highlights the discrepant messages, and directs people's attention to the deviant reviews. The nonsignificant results could be due to the test stimulus used. A unimodal distribution with some dispersion was used as the test stimulus in the study to represent a "dispersed distribution" due to its pervasion. For a unimodal distribution, the mode is very prominent, as there is only one peak. However, the dispersion of a unimodal distribution is not as prominent as the mode because the data points are clustered around the mode. Thus, it might not be sufficient to drive the effects. That is, due to a lack of prominence in the dispersion, the unimodal distribution does not effectively highlight the discrepant ratings. Therefore, the distribution presence does not direct people's attention to these deviant reviews.

In the evaluation stage, results show the expected effects of search time on product attitude. If consumers spend more time reading the consistent reviews, their product attitude and purchase intention become more in line with the provided average rating. If they spend more time reading the deviant ones, their product attitude and purchase intention become less in line with the average rating (towards the opposite direction). Besides, results suggest that the search variables fully mediate the search aids' effects on product attitude. These findings support prior research on the effect of search time on the outcome variables (Ratchford, Lee, and Talukdar 2006). In addition, the findings extend the prior research by investigating the influences of search aids as the antecedence of search time, and how search time mediates the effects of search aids on the outcome variables.

Across the two stages (search and evaluation), the presence of average rating only has a total negative effect on attitude under the condition that the distribution is absent and average rating is low. The distribution presence has a total negative effect on attitude, when the average rating is high. It also has a total positive effect, when the average rating is low. In sum, the total effects of average rating presence depend on the level of the average rating and the absence of a rating distribution. On the other hand, the presence of a distribution shifts the attitude in a direction opposite to the valence of the average rating.

Beside the quantitative analysis of Study 2, the analyses of the verbal reports from Study 1 and Study 2 also provide some additional insights. Results of

Study 1 confirm the prepurchase motive as the main motivator of online review searching behavior and the relevance of the two search aids, namely, the average rating and the rating distribution. Furthermore, results of Study 1 also show some preliminary support to the anchoring effect of the average rating presence and the prediction that the distribution presence diverts people's attention away from the average rating. These findings are largely in line with the quantitative analysis of Study 2 on the search process of consistent reviews. The qualitative analysis results of Study 2 confirm the prediction that the distribution presence leads to more search on the deviant reviews, although this effect is not found in the quantitative analysis of Study 2. In addition, the results of qualitative analysis of Study 2 also reveal some important considerations that people have during the search process. First, the qualitative results of study 2 show that consumers view consistent reviews or deviant reviews in a very broad sense. Most of them simply classify the reviews as "positive" or "negative" and consider all positive reviews as "consistent" with the average rating, when the average rating is high. This supports the operationalization of the classification between the consistent and deviant reviews for the quantitative analysis. Second, the qualitative results of study 2 confirm that the distribution could serve as a basis for judgment of the average rating, when the average rating is absent. This is due to the stimulus used in the current study (a unimodal distribution with some dispersion), in which the mode of the distribution is prominent and largely consistent with the average

rating. When the average rating is present, the presence of a distribution diverts people's attention away from the average rating. Such findings are in line with the quantitative results in Study 2, where the distribution presence does not affect preview ratio of the reviews when the average rating is absent due to the prominence and relevance of the mode, which cancels out the impact of the dispersion. The negative effect on consistent reviews driven by the dispersion can only be observed when the average rating is present. Third, the qualitative results of study 2 indicate that negative reviews are more diagnostic and unexpected, thus it attracts more attention, especially when the average rating is absent. Such findings are in line with the quantitative results about the evaluation stage in Study 2. Furthermore, the results are also largely consistent with prior studies that negative information is generally more diagnostic than positive information and is more useful for forming impressions (Skowronski and Carlston 1989). Lastly, to some extent, the results also provide some support for the assumption that in the absence of search aids, people search reviews by date or randomly as proposed in the baseline model.

Managerial Implications

Our findings that certain reviews in a discussion attract more attention and are processed to a greater extent than other reviews have several implications for product marketers. First, based on the results on the preview ratios on consistent reviews (see Figure 11 in results section), when the average rating is present and

the distribution is absent, the mean preview ratio of consistent reviews is the highest. When both search aids are present, the mean preview ratio of consistent reviews is the lowest. The preview ratio of consistent reviews, as a component of search time on consistent reviews, could cause consumers attitude and purchase intention to be more in line with the average rating. The classification of the reviews is valence-based in the current research. The results that consumers preview positive or negative reviews more under different conditions have important implications in how firms should respond to reviews selectively under different context. Prior research on corporate responses to customer feedbacks (both positive and negative) could be very beneficial. Responding to consumer complaints and accepting responsibility for a problem through justifications and apologies can produce greater customer satisfaction and willingness to do business with the company in the future (Conlon and Murray 1996; Tax, Brown, and Chandrashekar 1998). While negative reviews were to be handled seriously, firms are also increasingly paying attention to provide responses to positively reviews systematically (Park and Allen 2013). In this sense, the management views online reviews as a form of engagement with the consumers and building a long-term relationship in the virtual environment. Showing appreciation for the fact that a customer took the time to share his or her positive experience could help to build a courteous and responsive image. It could also help reinforce the positive comments and encourage future reviews. However, in practice, when a discussion

includes large amount of reviews, it costs the marketers much time and effort to respond to each message. Thus it is important for marketers to know which type of reviews to respond to more. Results suggest that when the average rating is low, marketers should prioritise the need of responding to more, if not all, consistent reviews (negative reviews in this case), under the condition that the average rating is present and the distribution is absent on the review forum. When the average rating is high, marketers should also reply to more consistent reviews (positive reviews in this case) to encourage positive posts, under the condition that the average rating is present and the distribution is absent.

Second, results show that average rating presence enhances the average search depth of both consistent and deviant reviews. Thus, when the average rating is present, marketers should read both the consistent and deviant reviews carefully and respond to them with dedicated effort. By responding in great detail to the negative reviews particularly, marketers can mitigate the undesirable impact of negative reviews or even impress the review readers. At the same time, by replying to the positive ones attentively, marketers can encourage more positive reviews in the future.

Third, online consumer review has become an important information source for market research and forecasting (Chen, Fay, and Wang 2003). They are free to access, provide up-to-the-minute information and reviews that are at the brand-model specific level. Online consumer review provides firms with important

information on consumer preference for both their competitors' and their own products. Given that the impact of search aids on the different reviews differs, marketers should place different weightages on various types of reviews as information source for market research and forecasting, when certain search aids are provided. Based on the results on the preview ratios on consistent reviews mentioned in the first point, when the average rating is present and the distribution is absent, the mean preview ratio of consistent reviews is the highest. When both search aids are present, the mean preview ratio of consistent reviews is the lowest. To improve the understanding of the online reviews' influence on consumers and to make an accurate market forecast, it is important for market researchers to know which type of reviews to collect and analyse. Results suggest that when the average rating is low, market researchers should collect more consistent reviews (negative reviews in this case), under the condition that the average rating is present and the distribution is absent on the review forum. When the average rating is high, marketers should also collect more consistent reviews (positive reviews in this case), under the condition that the average rating is present and the distribution is absent. In addition, results show that average rating presence enhances the average search depth of both consistent and deviant reviews. Thus, when the average rating is present, market researchers should analyze and interpret both consistent and deviant reviews carefully. By analyzing the review contents carefully, market researchers can have a better understanding of online reviews'

impact from the consumers' perspective, and develop a more accurate prediction for sales forecasting.

Given that search time on the reviews matters for consumers' product attitude and purchase intention, it is advisable for marketers to encourage consumers to read online reviews about their products, if their reviews are mostly positive. The marketers can make such encouragement in their marketing communications to consumers (e.g., advertisements and social media). Similarly, it is also advisable to encourage consumers to post reviews onto popular forums by providing them with incentives such as shopping vouchers and discounts on future purchases. For example, marketers can liaise with the popular review forums and identify each user's posting history. For the reviewers who have posted reviews on their products, marketers can email them certain incentives such as electronic coupons for their next purchase. In addition, location-based businesses such as retail operations can also take advantage of existing social mobile applications such as Foursquare. These businesses could declare their presence and populate their virtual space by posting various promotional offers and other marketing communication messages on Foursquare or the like. For example, a store owner could offer application users incentives (e.g., price discount) in order to encourage them to post tips about products or services of the store on the application platform. After posting their tips, users can claim the promotional offer at the store by flashing their posts on their mobile devices to the store employees.

The influences of search aids on search behaviors and consumer responses are also of special interests to the forum owners, as they provide practical guidelines for forum search aid designs. One of the main objectives of review forums is to facilitate consumers to acquire relevant information and evaluate the product under consideration. Overall, the SEM results show that the average rating presence and the distribution presence have significant effects on the search variables (e.g., preview ratio of consistent reviews) and/or attitude towards the focal product. In particular, the effects of the distribution presence on the attitude are significant regardless of the average rating level and the presence/absence of average rating. This suggests that consumers make use of the average rating statistic and the distribution in information gathering and product evaluation, thereby attesting to the value of these search aids. Thus, it is advisable for forum owners to provide these search aids, particularly the product rating distributions, in order to assist consumers in information search and product evaluation.

Limitations and Future Research

This paper has some limitations that offer opportunities for future research. First, based on the definitions of the search components, the data have some systematic missing values (the respondents that did not read both consistent and deviant reviews in detail). It is acknowledged that these missing values are not random because it is based on respondents' selections/choices and the data analyzed in this study is non-randomly selected or incidentally truncated (Greene

2011). Therefore, the findings may not be generalizable to the consumers who do not read either the consistent reviews or deviant reviews in detail. A multi-stage sample selection model (based on previewing and reading of different types of reviews) can be used to estimate the incidental truncation. Future research can focus on this aspect and use more rigorous methods to address the incidental truncation issue. Furthermore, future research can also focus on those people who do not preview certain type(s) of reviews (consistent or deviant) or read them in detail. The contextual factors (e.g., involvement in the purchase or time pressure) and/or personal characteristics (e.g., need for closure or gender) that lead to such search behaviors are worthy of further investigation. In addition, it would be useful to examine the decision-making strategy of consumers, such as how the search aids (average rating and rating distribution) the choice of their strategies. Also, future research can employ the Bayesian estimation approach to account for consumer heterogeneity in the estimation of model parameters.

Second, this research classifies reviews into consistent and deviant reviews based on the valence of a high or low average rating. However, such classification might not apply to a neutral average rating context (e.g., 3/5). With a neutral average rating, the classification of consistent and deviant reviews could be based on the extremity of the reviews instead of the valence. For example, the consumers may view the moderate reviews as consistent with a neutral average rating, and the extreme reviews as the inconsistent ones. Future research can investigate how

consumers categorize reviews under a neutral average rating context by conducting qualitative studies such as face-to-face interviews. This is worthy of further investigation as it can complement the current research model.

Third, this study looks into product-related consumer responses as the outcome variables, as they are of great interest to the product marketers. From the forum owner's perspective, it is also acknowledged that consumer review search behaviors can influence their responses to the review forum (e.g., search experience in the forum and revisiting intention) as well. A possible extension of the current research would therefore be to investigate how search behaviors influence forum-related consumer response. Of special interests to the forum managers is consumer experience at a review forum as the experience can greatly influence their attitudes towards the forum and revisiting intention. When making a decision or building product knowledge, if the consumers can access the most relevant reviews easily and experience much effort saving by using the forum search aids, they will form a positive attitude towards the forum and are more likely to visit the forum again in the future. This line of research is worthy of further investigation as they can contribute additional knowledge to the user online experience literature and provide practical guidelines for forum search aid designs. Fourth, the effects of presence of a distribution on certain search behaviors were not significant in our data. One possible explanation could be that the prominence in the dispersion of the test stimulus used might be insufficient to drive the effects.

The dispersion of a rating distribution reflects the heterogeneity of opinions. For a unimodal distribution, the data points are clustered around the mode. Therefore, a unimodal distribution might not effectively highlight the discrepant ratings and the distribution presence might not direct people's attention to these deviant reviews. To further investigate this issue, a bimodal distribution level can be included as the test stimuli for comparison in future study. A bimodal distribution has two modes spaced apart from each other. Therefore, the dispersion of a bimodal distribution is more prominent than that of a unimodal distribution and the effect proposed in the current model is expected to be found for a bimodal distribution. That is, the presence of a rating distribution can have a positive influence on the preview ratios of the deviant reviews, as it highlights the discrepant messages, and directs people's attention to these deviant reviews. At the same time, the distribution presence also can increase the average search depth of both the consistent and the deviant reviews. The presence of a product rating distribution with more prominent dispersion increases uncertainty, as it reflects certain level of disagreement. Thus, consumers are more motivated to allocate considerable cognitive effort during the search process in order to resolve the uncertainty.

Lastly, although the proposed model is developed under a general online review search context, it is tested under a specific beach resort setting. None of the individual differences, including need for closure, need for cognition, product knowledge, product involvement, prior online review experience was found to

affect the current research model. Nevertheless, it is acknowledged that some other contextual factors could affect the validity of the model. For example, research has found that product reviews have a greater effect on consumer search and purchase behavior for experience than for search goods (Huang et al. 2009; Senecal and Nantel 2004). Thus, it is expected that the effects of search time on product evaluation can be also moderated by the product type. Specifically, the effects of search time on product evaluation could be stronger for experience than for search goods. Thus, when extending the model to other product types, the findings should be interpreted with caution. In future research, various test products across different industries should be used in order to replicate the findings and establish the external validity of the model.

Appendix A Interview Guide for Study 1

Hello, I'm Guan Chong, a PhD candidate from the Nanyang Business School, NTU. I am currently doing research on how consumers respond to online consumer reviews. First of all, I would like to know if you have experiences searching online reviews before (Yes/ No). If the answer is yes, I will proceed with the interview.

I am interested in hearing about your experiences in this area, and would be asking you questions pertaining to the topic. All responses will be treated with utmost confidentiality. If at any point during the interview, you have any questions, feel free to stop me and ask me anything you want. Also, if there are questions that you do not wish to answer, let me know, and we will skip that question.

1. Based on your past experiences searching online reviews, tell me more about what motivate you to search online reviews.

2. What type of product do you usually search online reviews about?

3. (Showing the participant an example of online review discussion with multiple reviews about a product) Imagine if you are searching online reviews for a planned purchase in mind, how would you proceed to read them? How many reviews do you normally read under a discussion? Please tell me what you are looking at, thinking and feeling while performing the task.

4. Do you use the aggregated ratings (e.g., average ratings and distributions)?

How do you use them?

5. Do you use the review sorting functions (e.g., review with highest/lowest product rating first and most helpful review first)? After sorting the reviews, what will do next?

6. Under what circumstances will you more likely to be persuaded by the reviews?

We have come to the end of the interview. Do you have anything you want to add, to the answers that you gave earlier?

Thank you very much for your assistance in helping me to understand your experiences better. Once again, I want to assure you, that your confidentiality will be respected and you will not be identified in any way, in the final report.

Appendix B

Study 2 Guide

1. Arrive at least 30 minutes before the first session on each day.
2. Set up the experiment cubicles, mock review websites (Test 1-8) and instruction PPT slides page 1-2 (see Appendix B.1). Turn on Tobii Studio (online reviews project) and click on the test number of the next participant according to the assignment form. Open the Live View window in Tobii Studio. Turn the video camera on the wall away from the table (do not face the table). Make sure that the room is neatly arranged and there are no unnecessary materials inside (In case of Internet failure, please refer to the contingency plan in Appendix B.4; In the case of system failure or other emergencies, please refer to the contacts immediately in Appendix B.5).
3. Ensure enough copies of acknowledgment and consent forms, cash incentives, task pages (See Appendix B.2) and instruction pages (See Appendix B.3), an assignment form, an attendance sheet and stationery. Also, put up the signage “Online Consumer Behaviour Study. Dear participants: please kindly call or SMS the experimenter at xxxx xxxx when you have arrived” on the glass door at the lab entrance.
4. When a participant (P) arrives, ask her (or him) to present her matric or NRIC/FIN card and tell you her experiment time (cf. the timeslot for which she has signed up). Say:

“Hi good morning/afternoon, can I have your matric card or NRIC/FIN card, and the experiment time you have signed up for please?”

5. Check her name and time against the attendance sheet and the assignment form.

Put a tick beside her name if she is present, and a cross if she is not. If they match, ask P to read the consent form. Ask her to sign on the form if she is willing to participate voluntarily in the study. Then, ask her to switch off her handphone and lead her to the cubicle. Say:

“Please kindly switch off your handphone and follow me.”

If P comes too early, ask her to wait. Say:

“Sorry, please have a seat while waiting for your turn.”

If P comes at the wrong time, ask her to come again at the proper time. If her name does not appear on our schedule, ask her to register for a later time slot.

If she is more than seven minutes late, write down “late” beside her name and say “I will contact you again and arrange a make-up slot for you”.

6. Lead P to the room and ask her to put her personal belongings on a spare chair at the corner of the room and take a seat in front of the computer. Experiment starts. Play “recording one” from the computer.

Recording One:

“Hello, I’m Guan Chong, a PhD candidate from the Nanyang Business School, NTU. The experiment starts now and please kindly switch off your handphone. The current research is on how consumers respond to online

product reviews. Online product reviews are consumer comments about various products or services on review websites or forums. For example, popular product review websites include tripadvisor.com and Amazon.com. All responses will be treated with utmost confidentiality. If at any point during the interview, you have any questions, feel free to stop the experimenter and ask anything you want. Now, please read the instructions on “Slide One” on your computer screen and perform the task required”.

7. After P fills in the Stage 1 survey, give her a number puzzle as a filler task (See Appendix B.2), say “please read the instructions and perform the task required”.
8. After P finishes the filler task, bring P to the chair in front of the eye tracker, hand P the instruction page (See Appendix B.3), and say “please take a seat here, read the instructions and perform the task required”.
9. After P finishes writing the two questions, say “Please look at the screen and adjust your posture according to my instructions” and help the P to adjust until the two white dots are in the center of the screen. Say, “Please look at the blue dots as they appear on the screen and you may move your head naturally during the process”, then, calibrate the eye tracker by clicking on “start”.
10. After the calibration, if there is no missing calibration and the errors lines are short and largely within the circles, click on “accept” and name the recording with the corresponding P number and her test number on the assignment form,

e.g., P1_test1; if not, click on “recalibrate” and repeat the process. Say, “Please imagine that a school break is coming in two months time and you are searching online reviews about a beach resort that you may consider as your travel destination in this school break. Please start searching now. Also, please tell me when you finish viewing the above discussion by saying ‘FINISHED’. You may start once the website has loaded, and use your mouse as you wish to”.

11. Click on “start recording”. Show P the beach resort mock site corresponding to the test number assigned to her. Closely observe the P’s eye-movements (the red dots) on the Live View window. If the red dots disappear from the Live View, please assist P to adjust her posture until the dots appear again, e.g., ask her to move backwards (P’s eyes need to be about 60 cm away from the computer screen, and she needs to face directly to the screen).
12. After P finishes viewing, show the participant to the chair in front of the computer, say, “please take a seat here and click on Slide Two, read the instructions and perform the task required”.
13. After P finishes Stage 2 survey, click on “RTA” from the Tobii Studio and name the file with P number and test number. e.g.RTA_P1_test1, and say, “now I would like to ask you some questions pertaining to your search process.”¹
 - a1) What were you thinking of when you first approached the website?

a2) What ran through your head when you were searching the website? (Wait until P has finished answering this question; then ask the next one.) How did you conduct your information search? Can you describe your search process?

a3) How did you evaluate the resort using the information gathered?

a4)* Did you choose to search certain reviews based on the average product rating? Why/Why not? (If yes) How did you do it?

a5)* Did you choose to search certain reviews based on the product rating distribution? Why/Why not? (If yes) How did you do it?

a6) (If P used the sorting by rating function) Why did you use the sorting by rating function?

a7)* Did you search more reviews that are consistent with the average rating? Why?

a8)* Did you search more reviews that deviate from the average rating? Why?

¹ Note: ¹ Point 13 will be only conducted with 25% of the participants (randomly selected within each cell. *Only questions applicable to the corresponding P are asked. If a statement is unclear, say: "Could you elaborate further on this point?/Could you give me an example?/Could you give more details?/Could you be more specific?"

Stop the RTA by "F10" on the PC.

14. Experiment ends. Play recording two from the computer.

Recording Two:

"We have come to the end of the interview. Do you have anything you want to add to the answers that you gave earlier? Thank you very much for your assistance in helping me to understand your experiences better. Once

again, I want to assure you that your confidentiality will be respected and you will not be identified in any way in the final report.”

15. After P has completed the interview, write her name and matric number legibly on the acknowledgment form. Give him/her the six-dollar cash incentive and ask her to sign on the form. Say:

“Thank you very much for your cooperation. Here’s your cash incentive, please sign against your name on this acknowledgement form to certify you’ve received the money. Thank you for coming.”

Then, revert the PC and room to its original state. Wait for the next P to come.

16. At the end of the day, count the number of completed electronic audio and video files and check whether the number matches the number of participants indicated on the assignment list. Make sure that the cash balance matches the information on the acknowledgement form. Back up all the data and voice recorded electronic files of the day into a portable drive. Put the recorded electronic files in the same folder and number the folder according to P id.
17. Make sure that everything is neat and tidy before leaving the lab. Turn off the PCs and the rooms’ lights. Also, turn off other lights if you do not find anyone in the lab.

Study 2 Materials

Slide 1

Imagine that a school break is coming in two months' time and you are considering going to a beach resort for fun and relaxation in this school break. You have heard of a beach resort named Kamala Beach Resort from advertisements. The following contains a photo and a description of the Kamala Beach Resort from these advertisements.

'Kamala Beach Resort is one of the all-Inclusive resorts on the Phuket, Thailand. The resort has a variety of rooms, suites, cottages, casual and elegant restaurants, bars, a full service spa and water activities to enjoy. The resort is 35 minutes drive to Phuket Airport, 15 minutes drive to Patong, and 10 minutes walk to Kamala Village and Phuket Fantasea. Guest room rates range from SGD 70-150, depending on the room type and season'.

Bearing this in mind, please click on the hyperlink below and complete the following questionnaire. Please raise your hand when you finish.

[Survey 1 link](#)

Slide 2

Based on the above search process, please click on the hyperlink and complete the following questionnaire.

[Survey 2 link](#)

Task Page

Task: 6 by 6 Sudoku

Please fill a 6×6 grid with digits, so that each column, each row, and each of the nine 2×3 sub-grids that compose the grid (or block) contains all of the digits from 1 to 6.

4	2		6		1
			4		5
6	1				
				6	4
3		6			
2		1		4	6

Please raise your hand when you finish the puzzle.

Instruction Page

Remember that earlier I have asked you to imagine that a school break is coming in two months' time, you are considering going to a beach resort for fun and relaxation in this school break, you have heard of Kamala Beach Resort from advertisements and decided to know more about the resort from online reviews. You will be visiting a portion of a review web site devoted to Kamala Beach Resort to know more about the resort. Please visit the site as long as you want, just like what you do in real life when you search for information in a review forum. Before doing so, you should write two questions you have about this resort here.

Question

1 _____

Question

2 _____

Please raise your hand when you finish writing the two questions.

Samples Websites Based on the Manipulations

Samples websites based on the manipulation of average rating level (using average rating present and average rating present conditions as examples)

Average rating = high

Average rating = low

Resortreview.com

Kamala Beach Resort

Customer Reviews

181 reviews

Average Customer Review Rating
(181 Reviews)

3.7/10

Sorted By: Date▲ Rating

Page 1 2 3 19

"Disaster!"

Reviewed October 14, 2004 by JeremyCPortsmouth

Rating: 2/10

Without doubt the worst resort I have ever stayed in. My wife and I made the mistake of going here for our honeymoon and after 2 hours we checked out. The location of the resort is pretty bad being situated at the wrong end of the beach but it was the cleanliness (or lack of) that made this place so [\(Read More...\)](#)

"Dreadful service"

Reviewed January 4, 2006 by Amy S

Rating: 2/10

Myself and my girlfriend plus a friend stayed here after a delightful 4 day live aboard diving trip in the islands...and what a let down we had! The facilities and layout of the resort are quite alright but the staff and the service were some of the worst I have ever experienced. and I have worked 20 years in the travel trade and know my resorts. [\(Read More...\)](#)

"Just about passed muster"

Reviewed March 28, 2006 by William888

Rating: 3/10

Overall, this was a very quiet resort (only 30% occupied) which allowed us to leave the pool/beach/restaurant to ourselves. Excellent for winding down and a beautiful location. Negatives were: no easy access to a swimming beach as the tide went out 200+ meters, very bad for children due to lots of steps, furniture and fittings require urgent attention, potentially lethal [\(Read More...\)](#)

"Somewhat disappointed!"

Reviewed April 23, 2006 by anon

Rating: 3/10

After reading the reviews for the hotel I was expecting much more than what we ended up with. The hotel location had easy access to the beach and the grounds of the hotel were clean, however the surrounding area is still recovering from the tsunami. There's limited choice in Kamala for restaurants/bars, we found ourselves bored after the second night [\(Read More...\)](#)

"Dont stay here"

Reviewed April 23, 2006 by missdior

Rating: 2/10

As soon as we arrived here we hated it. Before we had even seen the room, we had decided to move the next day (My husband had booked it at the last minute for 5 days). The staff at the desk were surly and by Thai standards not welcoming at all. I had to double check it was the same [\(Read More...\)](#)

"What a dump!"

Reviewed April 27, 2006 by C K

Rating: 1/10

Had prebooked 4 nights in April, but my heart sank when I arrived. Travelling on my own, the Kamala offers loads of activities, which I was looking forward to delve into, but after only 2 hours on the premises, I had to get out and had arranged for a taxi to take me to another resort the following day. First [\(Read More...\)](#)

"Kamala Beach, trouble at paradise?"

Reviewed June 7, 2006 by Pimo

Rating: 3/10

188

Appendix C Survey Items

Prior Attitude Stage One (Bagozzi 1992)

7-point bipolar scales

“I ___ this resort”

dislike 1 2 3 4 5 6 7 like

unfavor 1 2 3 4 5 6 7 favor

disapprove 1 2 3 4 5 6 7 approve

“I think this resort is ___”

good 1 2 3 4 5 6 7 bad*

unpleasant 1 2 3 4 5 6 7 pleasant

Consumers’ Attitude Stage Two (Bagozzi 1992)

7-point bipolar scales

“I ___ this resort”

dislike 1 2 3 4 5 6 7 like*

unfavor 1 2 3 4 5 6 7 favor

disapprove 1 2 3 4 5 6 7 approve

“I think this resort is ___”

good 1 2 3 4 5 6 7 bad

unpleasant 1 2 3 4 5 6 7 pleasant

Evaluation Confidence (Berger and Mitchell 1989)

“I am ___ of my evaluations of the resort that I have just indicated”

certain 1 2 3 4 5 6 7 uncertain

not sure 1 2 3 4 5 6 7 sure

confident 1 2 3 4 5 6 7 not confident

Consumers’ Purchase Intention (Schlosser et al. 2006)

1 -Very Unlikely; 2- Unlikely; 3-Slightly Unlikely; 4-Neutral 5-Slightly Likely; 6-Likely; 7-Very likely

“I will visit this resort for holiday.”

Product Knowledge (Brucks 1985)

“Rate your knowledge of this product category, as compared to the average person.”

1-One of the LEAST knowledgeable; 2- Not knowledgeable; 3-Slightly knowledgeable; 4 -Somewhat knowledgeable; 5- knowledgeable; 6-Very knowledgeable; 7-One of the MOST knowledgeable

“Circle one of the numbers below to describe your familiarity with this product category.”

1-Not at all familiar; 2- Not familiar; 3-Slightly familiar; 4 -Somewhat familiar; 5-familiar; 6-Very familiar; 7-Extremely familiar

Involvement in the Product Category (Coulter, Price, and Feick 2003)

1 - Strongly disagree; 2 - Disagree; 3- Slightly disagree; 4 – Neutral; 5- Slightly agree; 6 - Agree; 7 -Strongly agree

The beach resorts that I choose to visit represent part of my self-image*

The beach resorts that I choose to visit portray an image of me to others*

The beach resorts that I choose to visit tell others about me*

Beach resorts are boring to me*

Beach resorts are fun to me

Beach resorts are fascinating to me

Beach resorts are important to me

Beach resorts are exciting to me*

The beach resorts that other people choose to visit tell me about them.

Need for Cognition Scale (Ailawadi, Neslin, and Gedenk 2001)

1 - Strongly disagree; 2 - Disagree; 3- Slightly disagree; 4 – Neutral; 5- Slightly agree; 6 - Agree; 7 -Strongly agree

“Thinking is not my idea of fun.”

“I like tasks that don’t require much thinking once I have learned them.”

“I only think as hard as I have to.”

Need for Closure (Roets and Hiel 2011)

1 - Completely disagree; 2 – Disagree; 3- Slightly disagree; 4 – Slightly agree; 5 - Agree; 6 - Completely agree

Preference for Predictability

I dislike unpredictable situations.

I dislike it when a person's statement could mean many different things.

I dislike questions which could be answered in many different ways.

I don't like to be with people who are capable of unexpected actions.

I feel irritated when one person disagrees with what everyone else in a group believes.*

Decisiveness

When I am confronted with a problem, I'm dying to reach a solution very quickly.

I would quickly become impatient and irritated if I would not find a solution to a problem immediately.

I do not usually consult many different opinions before forming my own view.*

When I have made a decision, I feel relieved.*

Online Review Experience (Arun & Shanker, 2011)

"I am a ____ in searching online product reviews for information"

a complete beginner 1 2 3 4 5 6 7 an expert

"I never search online reviews for information before 1 2 3 4 5 6 7 I have been searching online reviews for information for a long time."*

In the last 6 months, how often do you search online product reviews for information?

1-Everyday; 2- 4-6 times a week; 3- 1-3 times a week; 4- 1-3 times a month; 5- 3-5 times within 6-months; 6- 1-2 times within 6-months; 7-Not at all

Note. *These items were deleted after the confirmatory factor analysis.

Appendix D

Table 19
Unstandardized Coefficient Estimates of the Revised Joint Models Based on the Alternative Classifications of Consistent and Deviant Reviews

<i>Search Process</i>	Classification One^a		Classification Two^b		Classification Three^c	
	Low average rating	High average rating	Low average rating	High average rating	Low average rating	High average rating
Preview ratio of consistent reviews^d	R ² =.023 R ² _{adj.} =.011	R ² =.029 R ² _{adj.} =.017	R ² =.014 R ² _{adj.} =-.001	R ² =.017 R ² _{adj.} =.002	R ² =.008 R ² _{adj.} =-.004	R ² =.009 R ² _{adj.} =-.003
Average rating presence	0.158!^{e,f}	0.158!	0.084	0.084	0.124	0.124
Distribution presence	0.050	0.050	-0.098	-0.098	0.013	0.013
Average rating presence × Distribution presence	-0.291*	-0.291*	-0.058	-0.058	-0.162!	-0.162!
Preview ratio of deviant reviews	R ² =.002 R ² _{adj.} =-.002	R ² =.004 R ² _{adj.} =.000	R ² =.001 R ² _{adj.} =-.004	R ² =.002 R ² _{adj.} =-.003	R ² =.000 R ² _{adj.} =-.004	R ² =.000 R ² _{adj.} =-.004
Average rating presence	-0.071	-0.071	-0.048	-0.048	0.005	0.005
Average search depth of consistent reviews	R ² =.035 R ² _{adj.} =.027	R ² =.033 R ² _{adj.} =.025	R ² =.049 R ² _{adj.} =.040	R ² =.037 R ² _{adj.} =.028	R ² =.036 R ² _{adj.} =.028	R ² =.033 R ² _{adj.} =.025
Average rating presence	0.24*	0.24*	0.196!	0.196!	0.236*	0.236*
Distribution presence	-0.151	-0.151	-0.293**	-0.293**	-0.160!	-0.160!
Average search depth of deviant reviews	R ² =.020 R ² _{adj.} =.012	R ² =.032 R ² _{adj.} =.024	R ² =.016 R ² _{adj.} =.006	R ² =.032 R ² _{adj.} =.023	R ² =.016 R ² _{adj.} =.008	R ² =.027 R ² _{adj.} =.019
Average rating presence	0.257*	0.257*	0.235*	0.235*	0.239*	0.239*
Distribution presence	0.094	0.094	-0.013	-0.013	0.046	0.046

	Classification One		Classification Two		Classification Three	
<i>Search Process (Cont')</i>	Low average rating	High average rating	Low average rating	High average rating	Low average rating	High average rating
Search time on consistent reviews	R ² =.593 R ² _{adj.} =.590	R ² =.676 R ² _{adj.} =.673	R ² =.630 R ² _{adj.} =.623	R ² =.806 R ² _{adj.} =.802	R ² =.589 R ² _{adj.} =.582	R ² =.728 R ² _{adj.} =.724
Preview ratio of consistent reviews	0.577***	0.621***	0.712***	0.556***	0.662***	0.649***
Average search depth of consistent reviews	0.968***	1.093***	1.044***	1.003***	0.931***	1.147***
Search time on deviant reviews	R ² =.868 R ² _{adj.} =.867	R ² =.832 R ² _{adj.} =.829	R ² =.868 R ² _{adj.} =.866	R ² =.708 R ² _{adj.} =.702	R ² =.873 R ² _{adj.} =.871	R ² =.625 R ² _{adj.} =.619
Preview ratio of deviant reviews	0.799***	0.726***	0.087***	0.809***	0.811***	0.193*
Average search depth of deviant reviews	0.992***	0.967***	0.993***	0.994***	0.984***	0.997***
<i>Evaluation Process</i>						
Product attitude	R ² =.116 R ² _{adj.} =.100	R ² =.118 R ² _{adj.} =.104	R ² =.058 R ² _{adj.} =.041	R ² =.125 R ² _{adj.} =.106	R ² =.090 R ² _{adj.} =.074	R ² =.098 R ² _{adj.} =.083
Search time on consistent reviews	-0.326***	0.173**	-0.201**	0.063	-0.273***	0.129*
Search time on deviant reviews	0.173***	-0.326***	0.063	-0.201**	0.129*	-0.273***
Purchase intention	R ² =.796 R ² _{adj.} =.795	R ² =.823 R ² _{adj.} =.822	R ² =.737 R ² _{adj.} =.735	R ² =.757 R ² _{adj.} =.754	R ² =.748 R ² _{adj.} =.746	R ² =.808 R ² _{adj.} =.806
Product attitude	1.107***	1.107***	1.060***	1.060***	1.081***	1.081***

Note. ^a Under on the Classification One scheme, when the average rating is high, reviews with ratings 6 to 10 are classified as the consistent reviews and others are classified as the deviant reviews. When the average rating is low, reviews with rating 1 to 5 are classified as the consistent reviews and others are classified as the deviant reviews.

^b Under on the Classification Two scheme, when the average rating is high, reviews with ratings 7 to 8 are classified as the consistent reviews and others are classified as the deviant reviews. When the average rating is low, reviews with rating 3 to 4 are classified as the consistent reviews and others are classified as the deviant reviews.

^c Under on the Classification Three scheme, when the average rating is high, reviews with ratings 6 to 9 are classified as the consistent reviews and others are classified as the deviant reviews. When the average rating is low, reviews with rating 2 to 5 are classified as the consistent reviews and others are classified as the deviant reviews.

^d The variable in bold is regressed on the following variables.

^e Bold coefficients indicate support for hypothesis.

^f !p<.1, *p<.05, **p<.01, *** p<.001 (two-tailed).

Appendix E

Table 20

Unstandardized Estimates of the Bootstrap Results for Indirect Effects of Average Rating Presence and Distribution Presence

	Low average rating		High average rating	
	Distribution absent	Distribution present	Distribution absent	Distribution present
Effects from Average Rating Presence to Attitude				
<i>Total indirect effects</i>	-0.071* ^a (-0.154, -0.002) ^b	-0.016 (-0.087, 0.049)	-0.002 (-0.091, 0.084)	-0.033 (-0.117, 0.041)
<i>Specific indirect effects</i>				
average rating presence → preview ratio of consistent reviews → search time on consistent reviews → attitude	-0.030 (-0.070, 0.0010)	0.025 (-0.010, 0.069)	0.017 (-0.001, 0.044)	-0.014 (-0.043, 0.006)
average rating presence → average search depth of consistent reviews → search time on consistent reviews → attitude	-0.076* ^c (-0.156, -0.012)		0.045* (0.005, 0.103)	
average rating presence → preview ratio of deviant reviews → search time on deviant reviews → attitude	-0.010 (-0.038, 0.014)		0.017 (-0.023, 0.062)	
average rating presence → average search depth of deviant reviews → search time on deviant reviews → attitude	0.044* (0.006, 0.094)		-0.081* (-0.166, -0.015)	
	Low average rating		High average rating	
	Average Rating absent	Average Rating present	Average Rating absent	Average Rating present
Effects from Distribution Presence to Attitude				
<i>Total indirect effects</i>	0.054 (-0.004, 0.122)	0.109* (0.044, 0.192)	-0.053 (-0.117, 0.008)	-0.084* (-0.157, -0.013)
<i>Specific indirect effects</i>				
distribution presence → preview ratio of consistent reviews → search time on consistent reviews → attitude	-0.009 (-0.046, 0.025)	0.045* (0.013, 0.092)	0.005 (-0.014, 0.028)	-0.026* (-0.057, -0.005)
distribution presence → average search depth of consistent reviews → search time on consistent reviews → attitude	0.048 (-0.010, 0.120)		-0.029 (-0.074, 0.006)	
distribution presence → average search depth of deviant reviews → search time on deviant reviews → attitude	0.016 (-0.021, 0.060)		-0.030 (-0.098, 0.040)	

Note. ^a *p<.05 (two-tailed).

^b 95% confidence interval (two-tailed).

^c Only one coefficient estimate is presented when the effects are the same across the two conditions.

REFERENCE

- Ailawadi, Kusum L, Scott A Neslin, and Karen Gedenk (2001), "Pursuing the Value-Conscious Consumer: Store Brands Versus National Brand Promotions," *Journal of Marketing*, 65 (1), 71.
- Ajzen, I. (1988), *Attitudes, Personality, and Behavior*, Buckingham: Open University Press.
- Aribarg, Anocha, Rik Pieters, and Michel Wedel (2010), "Raising the Bar: Bias Adjustment of Recognition Tests in Advertising.," *Journal of Marketing Research*, 47 (3), 387-400.
- Ba, Sulin and Paul A. Pavlou (2002), "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly*, 26 (3), 243-68.
- Bagozzi, R. P. and T. F. Heatherton (1994), "A General Approach to Representing Multifaceted Personality Constructs: Application to State Self-Esteem," *Structural Equation Model*, 1 (1), 35-67.
- Bagozzi, R. and Y. Yi (2012), "Specification, Evaluation, and Interpretation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 40 (1), 8-34.
- Bagozzi, Richard P. (1992), "The Self-Regulation of Attitudes, Intentions, and Behavior," *Social Psychology Quarterly*, 55 (2), 178-204.
- Baron, J. (2000), *Thinking and Deciding (3rd Ed.)*, New York: Cambridge University Press.
- Bearden, W. O., R. G. Netemeyer, and M. F. Mobley (1993), *Handbook of Marketing Scales: Multi-Item Measures for Marketing and Consumer Behavior Research*, Newbury Park, CA: Sage Publications.
- Berger, I.E. and A.A. Mitchell (1989), "The Effect of Advertising on Attitude Accessibility, Attitude Confidence, and the Attitude-Behavior Relationship," *Journal of Consumer Research*, 16 (3), 269-79.

- Bettman, James R. and C. Whan Park (1980), "Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Analysis: A Protocol Analysis," *Journal of Consumer Behavior*, 7 (3), 234-48.
- Bickart, Barbara and Robert M. Schindler (2001), "Internet Forums as Influential Sources of Consumer Information," *Journal of Interactive Marketing*, 15 (3), 31-40.
- Bogdan, R. B. and S. K. Biklin (1998), *Qualitative Research for Education: An Introduction to Theory and Methods, Third Edition*, Needham Heights, MA: Allyn and Bacon.
- Bollen, Kenneth A. (1989), *Structural Equations with Latent Variables*, New York: Wiley.
- Brucks, Merrie (1985), "The Effects of Product Class Knowledge on Information Search Behavior," *Journal of Consumer Research*, 12 (1), 1-16.
- Cacioppo, John T., Richard E. Petty, Jeffrey A. Feinstein, W. Jarvis, and G. Blair (1996), "Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition," *Psychological Bulletin*, 119 (2), 197-253.
- Chapman, G.B. and E.J. Johnson (1999), "Anchoring, Activation and the Construction of Value," *Organizational Behavior and Human Decision Processes*, 79 (2), 1-39.
- Chatterjee, Patrali (2001), "Online Reviews: Do Consumers Use Them?" *Advances in Consumer Research*, 28, 129-33.
- Chen, Yubo, Scott Fay, and Qi Wang (2003), "Marketing Implications of Online Consumer Product Reviews," *Business Week*, 7150, 1-36.
- Chen, Yubo, Yong Liu, and Jurui Zhang (2011), "When Do Third-Party Product Reviews Affect Firm Value and What Can Firms Do? The Case of Media

- Critics and Professional Movie Reviews," *Journal of Marketing*, 74 (September), 116-34.
- Cheung, C., C. Luo, C. Sia, and H. Chen (2009), "Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of on-Line Consumer Recommendations," *International Journal of Electronic Commerce*, 13 (4), 9-38.
- Chevalier, J. A. and D. Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (3), 345–54.
- Clemons, E. K., G. Gao, and L. M. Hitt (2006), "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry," *Journal of Management Information Systems*, 23 (2), 149–71.
- Conlon, Donald E. and Noel M. Murray (1996), "Customer Perceptions of Corporate Responses to Product Complaints: The Role of Explanations," *Academy of Management Journal*, 39 (4), 1040-56.
- Coulter, Robin A. , Linda L. Price, and Lawrence F. Feick (2003), "Rethinking the Origins of Product Involvement and Brand Commitment: Insights from Postsocialist Central Europe," *Journal of Consumer Research*, 30 (September), 151-69.
- Craik, Fergus I.M. (2002), "Levels of Processing: Past, Present and Future?" *Memory*, 10 (5/6), 305-18.
- Davis, A. and D. Khazanchi (2008), "An Empirical Study of Online Word of Mouth as a Predictor for Multi-Product Category E-Commerce Sales," *Electronic Markets*, 18 (2), 130-41.
- Dellarocas, C., X. Zhang, and N. Awad (2007), "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing*, 21 (4), 23-45.

- Diehl, Kristin, L. J. Kornish, and J. G. Lynch (2003), "Smart Agents: When Lower Search Costs Forquality Information Increase Price Sensitivity," *Journal of Consumer Research*, 30 (1), 56-71.
- Duan, Wenjing, Bin Gu, and Andrew B. Whinston (2008), "Do Online Reviews Matter? - an Empirical Investigation of Panel Data," *Decision Support Systems*, 45 (4), 1007-16.
- Epley, N. and T. Gilovich (2005), "When Effortful Thinking Influences Judgmental Anchoring: Differential Effects of Forewarning and Incentives on Self-Generated and Externally-Provided Anchors.," *Journal of Behavioral Decision Making*, 18 (3), 199-212.
- Epley, Nicholoas and Thomas Gilovich (2006), "The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient," *Psychological Science*, 17(4), 311-18.
- Feldman, J. M. and J. G. Lynch (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (3), 421-35.
- Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Evanston, IL: Row, Peterson.
- Fiske, S. T. (1980), "Attention and Weight in Person Perception: The Impact of Negative and Extreme Behavior," *Journal of Personality and Social Psychology*, 38 (6), 889-906.
- Forman, C. , A. Ghose, and B. Wiesenfeld (2008), "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research*, 19 (3), 291-313.
- Fornell, C. and D. F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39-50.

- Freedman, Lauren (2008), *Merchant and Customer Perspective on Customer Reviews and User-Generated Content: The e-tailing group*.
- Gauri, D., A. Bhatnagar, and R. Rao (2008), "Role of Word of Mouth in Online Store Loyalty," *Communications of the ACM*, 51 (3), 89-91.
- Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545-60.
- Goldsmith, Ronald E. and David Horowitz (2006), "Measuring Motivations for Online Opinion Seeking," *Journal of Interactive Advertising*, 6 (2), 1-16.
- Greene, William (2011), *Econometric Analysis (7th Edition)* Prentice Hall.
- Gruen, T., T. Osmonbekov, and A. Czaplewski (2006), "Ewom: The Impact of Customer-to-Customer Online Know-How Exchange on Customer Value and Loyalty," *Journal of Business Research*, 59 (4), 449-56.
- Hair, J.F., R.E. Anderson, R.L. Tatham, and W.C. Black (1998), *Multivariate Data Analysis with Readings (5th Ed.)*, New York, NY: McMillan Publishing.
- Häubl, Gerald, Benedict G. C. Dellaert, and Bas Donkers (2010), "Tunnel Vision: Local Behavioral Influences on Consumer Decisions in Product Search," *Marketing Science*, 29 (3), 438–55.
- Haubl, Gerald and Valerie Trifts (2000), "Consumer Decision Making in Online Shopping. Environments: The Effect of Interactive Decision Aids," *Marketing Science*, 19 (1), 4 -21.
- Hennig-Thurau, Thorsten and Gianfranco Walsh (2003), "Electronic Word-of-Mouth: Motives for and Consequences of Reading Customer Articulations on the Internet," *International Journal of Electronic Commerce*, 8 (2), 51-74.
- Higgins, E. T. and W. Rholes (1978), "'Saying Is Believing': Effects of Message Modification of Memory and Liking for the Person Described," *Journal of Experimental Social Psychology*, 14, 363-78.

- Highhouse, S. and K. Bottrill (1995), "The Influence of Social (Mis)Information on Memory for Behavior in an Employment Interview," *Organizational Behavior and Human Decision Processes*, 62 (2), 220-29.
- Howard, J.A. (1989), *Consumer Behavior in Marketing Strategy*, Englewood Cliffs, NJ: Prentice Hall.
- Howard, J.A., R.P. Shay, and C.A. Green (1988), "Measuring the Effect of Marketing Information on Buying Intentions," *Journal of Services Marketing*, 2 (4), 27-36.
- Howard, J.A. and J.N. Sheth (1969), *The Theory of Buyer Behavior*, New York: Wiley.
- Hu, L. and P. M. Bentler (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives.," *Structural Equation Model*, 6 (1), 1-55.
- Hu, N., L. Liu, and J. Zhang (2008), "Do Online Reviews Affect Product Sales? The Role of Reviewer Characteristics and Temporal Effects," *Information Technology and Management*, 9 (3), 201-14.
- Hu, Nan, Jie Zhang, and Paul A. Pavlou (2009), "Overcoming the J-Shaped Distribution of Product Reviews," *Communications of the ACM*, 52 (10), 144-47.
- Huang, Peng, Nicholas H. Lurie, and Sabyasachi Mitra (2009), "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods," *Journal of Marketing*, 73 (2), 55-69.
- Huang, Yun Kuei and Wen I. Yang (2008), "Motives for and Consequences of Reading Internet Book Reviews," *The Electronic Library*, 26 (1), 97-110.
- Jacowitz, K.E. and D. Kahneman (1995), "Measures of Anchoring in Estimation Tasks," *Personality and Social Psychology Bulletin*, 21 (11), 1161-66.
- Jonas, E., S. Schulz-Hardt, and D. Frey (2005), "Giving Advice or Making Decisions in Someone Else's Place: The Influence of Impression, Defense

- and Accuracy Motivation on the Search for New Information," *Personality and Social Psychology Bulletin*, 31 (7), 977-90.
- Juster, T. F. (1966), "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design," *Journal of the American Statistical Association*, 61 (315), 658-96.
- Kenny, David (1979), *Correlation and Causality*, New York: Wiley.
- Kraus, SJ (1995), "Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature," *Personality and Social Psychology Bulletin*, 21 (1), 58-75.
- Kruglanski, A. W and T. Freund (1983), "The Freezing and Unfreezing of Lay Inferences: Effects of Impressional Primacy, Ethnic Stereotyping, and Numerical Anchoring," *Journal of Experimental Social Psychology*, 19 (5), 448-68.
- Kruglanski, A.W., A. Pierro, L. Mannetti, and E. De Grada (2006), "Groups as Epistemic Providers: Need for Closure and the Unfolding of Group-Centrism," *Psychological Review*, 113 (1), 84-100.
- Kunda, Z. (1999), *Social Cognition*, Cambridge, Mass: MIT Press.
- Lam, Shun Yin, Albert Wai-Lap Chau, and Tsunhin John Wong (2007), "Thumbnails as Online Product Displays: How Consumers Process Them," *Journal of Interactive Marketing*, 21 (1), 36-59.
- Laroche, M., C. Kim, and L. Zhou (1996), "Brand Familiarity and Confidence as Determinants of Purchase Intention: An Empirical Test in a Multiple Brand Context," *Journal of Business Research*, 37 (2), 115-20.
- Laroche, M. and R. Sadokierski (1994), "Role of Confidence in a Multi-Brand Model of Intentions for a High-Involvement Service," *Journal of Business Research*, 29 (1), 1-12.

- Lee, J., D. Park, and I. Han (2008), "The Effect of Negative Online Consumer Reviews on Product Attitude: An Information Processing View," *Electronic Commerce Research and Applications*, 7 (3), 341-52.
- Lilien, Gary L., Philip Kotler, and K. Sridhar Moorthy (1995), *Marketing Models*, Englewood Cliffs, NJ. : Prentice Hall.
- Lingle, John H., Nehemia Geva, Thomas M. Ostrom, Michael R. Leippe, and Michael H. Baumgartner (1979), "Thematic Effects of Person Judgments on Impression Organization," *Journal of Personality and Social Psychology*, 37 (5), 674-87.
- Liu, Yong (2006), "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70 (3), 74 - 89.
- Matz, David C. and Wendy Wood (2005), "Cognitive Dissonance in Groups: The Consequences of Disagreement," *Journal of Personality and Social Psychology*, 88 (1), 22-37.
- McCracken, G. (1988), *The Long Interview*, Newberry Park, CA: Sage.
- Menon, Geeta, Priya Raghuram, and Norbert Schwarz (1995), "Behavioral Frequency Judgments: An Accessibility-Diagnosticity Framework," *Journal of Consumer Research*, 22 (2), 212-28.
- Moe, Wendy W. (2003), "Buying, Searching, or Browsing: Differentiating between Online Shoppers Using in-Store Navigational Clickstream," *Journal of Consumer Psychology*, 13 (1&2), 29-39.
- Moneta, G. B. and P. P. Yip (2004), "Construct Validity of the Scores of the Chinese Version of the Need for Closure Scale," *Educational & Psychological Measurement*, 64 (3), 531-48.
- Moorthy, Sridhar, Brian T. Ratchford, and Debabrata Talukdar (1997), "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Research*, 23 (4), 263-77.

- Moskowitz, G. B. (1996), "The Mediational Effects of Attributions and Information Processing in Minority Social Influence," *British Journal of Social Psychology Quarterly*, 35 (1), 47-66.
- Murray, Kyle B. and Gerald Haubl (2008), "Interactive Consumer Decision Aids," in *Handbook of Marketing Decision Models*, Vol. 121, LLC. NY., 55-77.
- Muthén, L. K. and B. O. Muthén (2011), *Mplus User's Guide. Sixth Edition*, Los Angeles, CA: Muthén & Muthén.
- Neuberg, S. L., S. G. West, T. N. Judice, and M. M. Thompson (1997), "On Dimensionality, Discriminant Validity, and the Role of Psychometric Analyses in Personality Theory and Measurement: Reply to Kruglanski et al.'s (1997) Defense of the Need for Closure Scale," *Journal of Personality and Social Psychology*, 73 (5), 1017-29.
- Nickerson, Raymond S. (1998), "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.," *Review of General Psychology*, 2 (2), 175-220.
- Park, C. and T. Lee (2009), "Information Direction, Website Reputation and Ewom Effect: A Moderating Role of Product Type," *Journal of Business Research*, 62 (1), 61-67.
- Park, Do-Hyung, Jumin Lee, and Ingoo Han (2007), "The Effect of on-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement," *International Journal of Electronic Commerce*, 11 (4), 125-48.
- Park, Sun-Young and Jonathan P. Allen (2013), "Responding to Online Reviews : Problem Solving and Engagement in Hotels," *Cornell Hospitality Quarterly*, 54 (1), 64-73.
- Patton, M. Q. (1990), *Qualitative Evaluation and Research Methods(2nd Ed.)*, Newbury Park, CA: Sage.

- Payne, J.W., J.R. Bettman, and E.J. Johnson (1988), "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (3), 534-52.
- Payne, John W. and James R. Bettman (1992), "Behavioral Decision Research: A Constructive Processing Perspective," *Annual Review of Psychology*, 43 (1), 87-131.
- Petkova, KG, I Ajzen, and BL. Driver (1995), "Salience of Anti-Abortion Beliefs and Commitment to an Attitudinal Position: On the Strength, Structure, and Predictive Validity of Anti-Abortion Attitudes," *Journal of Applied Social Psychology*, 25 (6), 463-83.
- Petty, RE and JA Krosnick (1995), *Attitude Strength: Antecedents and Consequences.*, Mahwah, NJ: Erlbaum.
- Petty, Richard E. and John T. Cacioppo (1996), *Attitudes and Persuasion: Classic and Contemporary Approaches*, Boulder, CO: Westview Press.
- Pieters, Rik (2008), "A Review of Eye-Tracking Research in Marketing," in *Review of Marketing Research*, Vol. 4, ed. Naresh K. Malhotra, Armonk, NY: Emerald Group Publishing Limited, 123-47.
- Pieters, Rik and Luk Warlop (1999), "Visual Attention During Brand Choice: The Impact of Time Pressure and Task Motivation," *International Journal of Research in Marketing*, 16, 1-17.
- Ping, Robert A Jr. (2004), "On Assuring Valid Measures for Theoretical Models Using Survey Data," *Journal of Business Research*, 57 (2), 125-41.
- Ping, Robert A. (1995), "A Parsimonious Estimating Technique for Interaction and Quadratic Latent Variables," *Journal of Marketing Research*, 32 (3), 336-47.
- Putrevu, S. and B. Ratchford (1997), "A Model of Search Behavior with an Application to Grocery Shopping," *Journal of Retailing*, 73 (4), 463-86.

- Ratchford, Brian T. (2001), "The Economics of Consumer Knowledge," *Journal of Consumer Research*, 27 (4), 397-411.
- Ratchford, Brian T., Myung-Soo Lee, and Debabrata Talukdar (2006), "Consumer Use of the Internet in Search for Automobiles: Literature Review, a Conceptual Framework, and an Empirical Investigation," in *Review of Marketing Research*, Vol. 2, ed. Naresh K. Malhotra: Emerald Group Publishing Limited, 81-108.
- Ratchford, Brian T. and Narasimhan Srinivasan (1993), "An Empirical Investigation of Returns to Search," *Marketing Science*, 12 (1), 73-87.
- Rayner, Keith (1998), "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological Bulletin*, 124 (3), 372-422.
- Rayner, Keith (2009), "Eye Movements and Attention in Reading, Scene Perception, and Visual Search," *The Quarterly Journal of Experimental Psychology*, 62 (8), 1457-506.
- Roets, Arne and Alain Van Hiel (2011), "Item Selection and Validation of a Brief, 15-Item Version of the Need for Closure Scale," *Personality and Individual Differences*, 50 (1), 90-94.
- Roets, Arne and Alain Van Hiel (2011), "Item Selection and Validation of a Brief, 15-Item Version of the Need for Closure Scale," *Personality and Individual Differences*, 50 (1), 90-94.
- Rogers, E. M. (2003), *Diffusion of Innovations (5th Ed.)*, New York, NY: Free Press.
- Ross, Ben (2010), *Inside Travel: Tripadvisor: The independent*.
- Russo, J. Edward (1978), "Eye Fixations Can Save the World: A Critical Evaluation and a Comparison between Eye Fixations and Other Information Processing Methodologies," in *Advances in Consumer Research*, ed. H.K. Hunt, Ann Arbor, MI: Association for Consumer Research, 561-70.

- Russo, J. Edward (2011), "Eye Fixations as a Process Trace," in *A Handbook of Process Tracing Methods for Decision Research : A Critical Review and User's Guide*, ed. Michael Schulte-Mecklenbeck, Anton Kühberger and Rob Ranyard, New York: Psychology Press.
- Russo, J. Edward and France Leclerc (1994), "An Eye-Fixation Analysis of Choice Processes for Consumer Nondurables," *Journal of Consumer Research*, 21 (2), 274-90.
- Rust, R. T. and B. Cooil (1994), "Reliability Measures for Qualitative Data: Theory and Implications," *Journal of Marketing Research*, 31 (1), 1-14.
- Sauer, Paul L. and Alan Dick (1993), "Using Moderator Variables in Structural Equation Models," in *Advances in Consumer Research*, Vol. 20, ed. Leigh McAlister and Michael L. Rothschild: Association for Consumer Research, 636-40.
- Schlosser, Ann E. (2003), "Experiencing Products in the Virtual World: The Role of Goal and Imagery in Influencing Attitudes Versus Purchase Intentions," *Journal of Consumer Research*, 30 (September), 184–98.
- Schlosser, Ann E., Tiffany Barnett White, and Susan M. Lloyd (2006), "Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions," *Journal of Marketing*, 70 (April), 133-48.
- Schmidt, Jeffrey B. and Richard A. Spreng (1996), "A Proposed Model of External Consumer Information Search," *Journal of the Academy of Marketing Science*, 24 (3), 246-56.
- Senecal, Sylvain and Jacques Nantel (2004), "The Influence of Online Product Recommendations on Consumers' Online Choices," *Journal of Retailing*, 80 (2), 159-69.

- Sher, Peter J. and Sheng-Hsien Lee (2009), "Consumer Skepticism and Online Reviews: An Elaboration Likelihood Model Perspective," *Social Behavior and Personality*, 37 (1), 137.
- Simmons, Carolyn J., Barbara Bickart, and John G. Lynch (1993), "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 20 (2), 316-29.
- Skowronski, John J. and D. E. Carlston (1989), "Negativity and Extremity Biases in Impression Formation: A Review of Explanations," *Psychological Bulletin*, 105 (1), 131-42.
- Stalder, D. R. (2007), "Need for Closure, the Big Five, and Public Self-Consciousness," *Journal of Social Psychology*, 147 (1), 91-94.
- Stalder, D. R. (2009), "Competing Roles for the Subfactors of Need for Closure in Committing the Fundamental Attribution Error," *Personality and Individual Differences*, 47 (7), 701-05.
- Steffes, Erin M. and Lawrence E. Burgee (2009), "Social Ties and Online Word of Mouth," *Internet Research*, 19 (1), 42.
- Stigler, George (1961), "The Economics of Information," *Journal of Political Economy*, 69 (3), 213-25.
- Tabachnick, Barbara G. and Linda S. Fidell (2007), *Using Multivariate Statistics, 5th Ed.*, New York: Allyn and Bacon.
- Tax, Stephen S., Stephen W. Brown, and Murali Chandrashekar (1998), "Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing," *Journal of Marketing*, 62 (April), 60-76.
- Tormala, Z.L., J.J. Clarkson, and R.E. Petty (2006), "Resisting Persuasion by the Skin of One's Teeth: The Hidden Success of Resisted Persuasive Messages.," *Journal of Personality and Social Psychology Bulletin*, 91 (3), 423-35.

- Treisman, A.M. (1979), "The Psychological Reality of Levels of Processing," in *Levels of Processing in Human Memory*, ed. L.S. Cermak & F.I.M. Craik, Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 301-30.
- Trusov, Michael, Randolph E. Bucklin, and Koen Pauwels (2009), "Modelling Effects of Word-of-Mouth Versus Traditional Marketing Findings from an Internet Social Networking Site," *Journal of Marketing*, 73 (5), 90-102.
- Tversky, A. and D. Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185 (4157), 1124-31.
- Tversky, A. and D. Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (4), 297-323.
- Van Hiel, A., M. Pandelaere, and B. Duriez (2004), "The Impact of Need for Closure on Conservative Beliefs and Racism: Differential Mediation by Authoritarian Submission and Authoritarian Dominance," *Personality & Social Psychology Bulletin*, 30 (7), 824-37.
- Verplanken, Bas, Pieter T Hazenberg, and Grace R Palenwen (1992), "Need for Cognition and External Information Search Effort," *Journal of Research in Personality*, 26 (2), 128-36.
- Wason, P. C. (1968), "Reasoning About a Rule," *Quarterly Journal of Experimental Psychology*, 20 (3), 271-81.
- Webster, D. and A. Kruglanski (1994), "Individual Differences in Need for Cognitive Closure," *Journal of Personality and Social Psychology*, 67 (6), 1049-62.
- Wedel, Michel and Rik Pieters (2000), "Eye Fixations on Advertisements and Memory for Brands: A Model and Findings," *Marketing Science*, 19 (4), 297-312.

- West, P.M. and S.M. Broniarczyk (1998), "Integrating Multiple Opinions: The Role of Aspiration Level on Consumer Response to Critic Consensus," *Journal of Consumer Research*, 25 (4), 38-51.
- Wright, Malcolm and Murray MacRae (2007), "Bias and Variability in Purchase Intention Scales," *Journal of the Academy of Marketing Science*, 35 (4), 617-24.
- Xia, L. and N. Bechwati (2008), "Word of Mouse: The Role of Cognitive Personalization in Online Consumer Reviews," *Journal of Interactive Marketing*, 9 (1), 3-13.
- Zaichkowsky, J. L. (1985), "Measuring the Involvement Construct," *Journal of Consumer Research*, 12 (3), 341-52.
- Zemborain, M. R., & Johar, G. V. (2007), "Attitudinal Ambivalence and Openness to Persuasion: A Framework for Interpersonal Influence," *Journal of Consumer Research*, 33 (4), 506-14.
- Zhang, W. and S. A. Watts (2008), "Capitalizing on Content: Information Adoption in Two Online Communities," *Journal of the Association for Information Systems*, 9 (2), 73-94.
- Zhao, Xinshu, Jr. John G. Lynch, and Qimei Chen (2010), "Reconsidering Baron and Kenny: Myths and Truths About Mediation Analysis," *Journal of Consumer Research*, 37 (August), 197-206.
- Zhu, Feng and Xiaoquan (Michael) Zhang (2010), "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, 74 (2), 133-48.
- Zwick, R., A. Rapoport, A. K. C. Lo, and A. V. Muthukrishnan (2003), "Consumer Sequential Search: Not Enough or Too Much?," *Marketing Science*, 22 (4), 503-19.