Spring 2018

Numerical Framing and Emotional Arousal as Moderators of Review Valence and Consumer Choices

Anh Dang
Old Dominion University

NUMERICAL FRAMING AND EMOTIONAL AROUSAL AS MODERATORS OF REVIEW VALENCE AND CONSUMER CHOICES

By

Anh Dang

B.B.A October 2010, International University, Vietnam

M.B.A May 2013, University of Houston – Clear Lake, Texas, USA

A Dissertation Submitted to the Faculty of

Old Dominion University in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

BUSINESS ADMINISTRATION-MARKETING

(MINOR IN INTERNATIONAL BUSINESS)

OLD DOMINION UNIVERSITY

May 2018

Approved by:

Yuping Liu-Thompkins (Director)

Mahesh Gopinath (Member)

Wu He (Member)

ABSTRACT

NUMERICAL FRAMING AND EMOTIONAL AROUSAL AS MODERATORS OF REVIEW VALENCE AND CONSUMER CHOICES

Anh Dang

Old Dominion University, 2018

Director: Yuping Liu-Thompkins

Online reviews are gaining importance in determining consumers' purchase decisions since many consumers trust them as much as personal word-of-mouth. One aspect of reviews that has received great research attention is valence. Valence refers to consumers' positive or negative evaluations of products. It can be reflected by star ratings or dichotomous choices such as recommendation rates and thumbs up or down rates. The effects of valence reported in previous studies have been equivocal at best. Therefore, the purpose of this dissertation is to identify factors that help reconcile these inconclusive findings.

The first essay examined emotional arousal (e.g., sad versus angry) as a moderator of the relationship between valence and consumer decisions. Through two lab experiments and one field study utilizing the browsing and purchasing data from a major online retailer, I find that the effect of emotional arousal can be different along the consumer purchase journey. During the search stage, consumers use emotional arousal as a heuristic to make their choices. Extreme reviews (e.g., five-star or one-star rating) with high emotional arousal indicate reviewers' bias and lack of self-

control and are deemed less informative about product performance. Therefore, emotional arousal weakens the effect of valence on consumers' consideration choices. However, when consumers are at the purchase stage, a more complex cognitive process emerges. Even though they believe that extremely negative reviews with high emotional arousal are uninformative, their anticipated regret leads them to reject products associated with those reviews.

The second essay suggests that how consumers process valence and volume (i.e., the total number of a product's reviews) depends on the framing of the numeric information, which subsequently determines the importance of valence in relation to that of volume in consumers' purchase decisions. Specifically, consumers will utilize different approaches to processing valence and volume information when valence is framed as a percentage of volume (60% of 500 customers recommend) versus when it is represented as an absolute number (e.g., 300 out of 500 customers recommend). Through five lab experiments (including an eye-tracking study), I find that due to the fundamental differences between these approaches, consumers are likely to tradeoff valence for high volume if the valence information is expressed as percentages. However, the dominant effect of review volume diminishes if the absolute number format is applied. The effect of numerical framing thus helps newly introduced high-quality products overcome their disadvantage due to low review volume.

I dedicate this dissertation to my parents, my parents-in-law, my brother and sister-in-law, and my husband. Their unconditional love and support inspires me to strive for successes.

ACKNOWLEDGEMENT

I encountered many challenges while completing my doctoral degree. Without the great support and help from the following precious people, my accomplishment would have been impossible.

I would first like to express my gratitude to Dr. Yuping Liu-Thompkins. Being her student is one of the greatest experiences I have ever had in my life. As my advisor, she inspired me to love both research and teaching. Whenever a research or a teaching problem arose, I always got great advice from her to overcome it. Doctoral students usually find it hard to finish their dissertations on time. Yet, I could still finish mine as planned thanks to Dr. Liu-Thompkins. She always made sure that we met regularly; thus, I never felt slacking. In addition, she equipped me with productive and up-to-date research tools that facilitated my dissertation well. I would also like to thank my committee members, Dr. Mahesh Gopinath and Dr. Wu He. They provided plenty of insightful comments for me to improve my dissertation.

A very special thanks to Dr. John B. Ford for the knowledge he conveyed as well as the immense support he provided me as the director of the Ph.D. in Business Administration program. When I presented my first conference paper, I made many mistakes but Dr. Ford was very encouraging. Similarly, in class, for any of mine papers, his comments were always supportive. Thanks to his encouragement, I never have doubt about the research and teaching activities that I have chosen for my career.

I am grateful for Dr. Aaron Arndt with whom I had my first publication. I learned from him not only how to do research but also his positive attitude towards work. My project with him helped prepare me well for the dissertation. I also learned a lot from Dr. Kiran Karande. It was very enjoyable to communicate and work with him.

I would like to extend my thanks to Ms. Katrina J. Davenport. She instructed me in details how to handle administrative work. In addition, she always reminded me of important events that are beneficial for doctoral students.

In addition, my friends played an important role in the accomplishment of my dissertation. My cohort – Olga Butenko-Moseley and Hangjun Xu, and I worked together to overcome the hardest times of the program. Likewise, we shared happiness moments too. Thanks to Khiem Huynh, a dear Vietnamese friend of mine, my talk with him always helped me keep a positive view of challenges.

I am deeply indebted to my parents, my parents-in-law, my brother and sister-in-law, and my awesome husband. Talking to my parents and my in-laws reinforces my commitment to the program. They were not pushy; rather, they were very understanding. The stories they told me about their youth really inspired me. I feel blessed to have Trung Nguyen as my husband. Words cannot describe how great he is. He is very smart, caring, and understanding. He has always been behind me regardless of his busy schedule. The picture of my doctoral study time would be incomplete without him. He was the most beautiful puzzle in that picture.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

ESSAY 1

DOES EMOTIONAL AROUSAL BOOST OR DISCOUNT REVIEW VALENCE? THE EFFECT OF PURCHASE STAGE

ABSTRACT

Extant research has pointed to the influence of review star ratings on consumers' purchase decisions. However, conflicting findings exist regarding the impact of star ratings. The present paper aims to reconcile these findings by examining emotional arousal as a moderator of the relationship between star ratings and consumer decisions. Through two lab experiments and one field study, this essay shows that emotional arousal can play two opposing roles depending on which purchase stage the consumer is at. During the search stage, it can prevent consumers from relying on valence since consumers perceive reviews with high arousal to be less informative about product performance and informativeness determines their consideration choices. Meanwhile, when consumers are about to make their purchase decision, negativity bias causes consumers to reject products with negatively high arousal reviews, even though they still believe that negative reviews with higher arousal are less informative. These two differential effects of emotional arousal were demonstrated along consumer purchase journey. The findings can help guide more effective leverage of consumer reviews, and offer a more dynamic and personalized view of the review management process.

INTRODUCTION

Word-of-mouth (WOM) plays an important role in consumers' decision making since it helps them reduce search costs as well as purchase risks (King et al. 2014). Thanks to the advance in information technology, consumers can receive word-of-mouth from many others through online review platforms. The problem that consumers face nowadays is no longer the lack of word-of-mouth, but rather their uncertainty of the quality of WOM expressed in online reviews (Ludwig et al. 2013, King et al. 2014). Accordingly, despite the existence of online reviews on their retail site, firms still have difficulty influencing consumer purchase decisions (Bonnet and Nandan 2011). To address this major concern, prior research has spent a great deal of effort in uncovering the factors that enhance the helpfulness of online reviews. One among them is review valence that refers to consumers' (positive or negative) numerical indication of product performance and their experience (e.g., one star rating) (Mudambi and Schuff 2010, Chevalier and Mayzlin 2006). Even though review valence has been considered a factor driving product sales in many studies of online word of mouth (e.g., Chevalier and Mayzlin 2006, Li and Hitt 2008), the effects of valence are still equivocal (King et al. 2014).

The inconclusive findings of prior research calls for the examination of other factors beyond review valence. Whereas valence can be informative to consumers, the limited choice of valence (e.g., 1 star to 5 star ratings) restrains the amount of word-of-mouth that reviewers can share. Therefore, prior studies have not only investigated valence but also review content (e.g., Schlosser 2011, Chen and Lurie 2013, Zou et al. 2011). For the same level of valence, two consumers can show two distinct product experiences in their written reviews. For example, while one consumer may feel "satisfied" with a product experience, another may describe the experience with “excitement”. Similarly, one consumer with negative product experiences may express “disappointment”, while another may feel “angry” or “mad”. Emotional arousal, i.e., the level of energy associated with an emotional experience (Niedenthal 2008, Russell 1980), can signal important information about those who express it such as their deliberation and self-control (Parrott 1995, Yin et al. 2017). The lack of self-control signified by high emotional arousal then can cause consumers to discount the information contained in the review. Therefore, it is likely that consumers use emotional arousal to adjust the information provided by the star rating. Previous studies show that emotional can influence consumers’ use of reviews. For example, Yin et al. (2017) report the direct effect of emotional arousal on review helpfulness. Ordenes et al. (2017) also demonstrated the importance of emotional arousal by including it as a component of sentiment strength that subsequently affects future product sales. These findings allude to the potential effect of emotional arousal on consumer purchases. However, the analysis was conducted at the aggregate level, which may inaccurately reflect individual level reactions. In addition, these previous studies have examined the main effect of emotional arousal in parallel with other review factors. Yet it is possible for emotional arousal to interact with other review factors such as valence on consumer purchase decisions. The current research examines this possibility.

Extant literature has shown that the more extreme the valence of a review is (e.g., five-star rating versus four-star ratings, one-star vs. two-star ratings), the more influential the review is in consumers’ decisions (Pavlou and Dimoka 2006, Forman et al. 2008). My research argues that emotional arousal in a review can moderate the influence of valence extremity on consumer purchase decisions. In this capacity, emotional arousal can play two opposing roles. On one hand, consumers may not follow extreme reviews with high emotional arousal since they perceive these reviews to be biased and not indicative of true product performance (Bodenhausen et al. 1994, Linville & Jones 1980, Linville 1982). Therefore, extremely negative reviews with high arousal (vs. low arousal) make the products deemed less negative and those high arousal positive reviews cause the products to look less positive. When consumers select products for their consideration set during their search stage, they prefer products with high arousal one-star reviews to those with two-star reviews. Similarly, they avoid products with high arousal five-star reviews and select those with four-star reviews. On the other hand, consumer use of emotional arousal is different when they are at their purchase stage and ready to make their decision. Specifically, whereas consumers at the search stage still have a second chance to select a product from their consideration set, consumers at the purchase stage are more risk-averse since they do not want to have any regret after buying a product. Therefore, for negative reviews, consumers tend to avoid a product with high arousal one-star reviews in case such an arousal feeling expressed in those reviews occurs to them if they buy that product. Accordingly, these two roles of emotional arousal arise in different stages along the consumer journey. This stems from previous research that shows the decision making criteria that consumers employ in their search stage can be different from those used in their purchase stage (Bettman and Park 1980). Hence, emotional arousal in online reviews can play two opposing roles depending on which stage the consumer belongs to. These effects were tested through two lab experiments and one field study.

My research provides several important contributions to marketing research and practice. First, by examining the interaction between star ratings and emotional arousal, my studies help reconcile the inconsistent findings from prior research regarding the role of star ratings. Specifically, even though star ratings signify product performance, consumers will not always utilize the information implied by valence depending on the emotional arousal reflected in the reviews. Second, by studying both the search and the purchase stage, the paper demonstrates that consumers' use of reviews depends on where they are in the purchase journey. Focusing exclusively on the final purchase stage and overlooking the search stage may induce researchers to draw misleading conclusions on the effect of review valence as well as emotional arousal. Finally, by leveraging real-world data on consumers' actual review reading and purchase behaviors captured by a major online retailer, I am able to test directly the impact of star ratings on individual consumer decisions and can provide more directly applicable insights to businesses.

LITERATURE REVIEW: MIXED EFFECTS OF REVIEW VALENCE

Prior studies suggest that review valence can influence consumers' purchases. For instance, Ye et al. (2010) show that a 10 percent increase in traveler valence can enhance online bookings by more than 5 percent. Furthermore, some papers argue for the positivity bias as they show that consumers rely more on positive reviews than they do on negative reviews (e.g., Li and Hitt 2008). Such bias is due to consumers' tendency to confirm their already purchased products (King et al. 2014). Meanwhile, other studies reveal the presence of negativity bias. For instance, Chevalier and Mayzlin (2006) found that one-star reviews discourage sales to a greater extent than five-star reviews because the latter seems less credible to consumers. In contrast, Berger and colleagues (2010) posit that negative reviews can increase sales by increasing product awareness. Briefly, the impact of review valence is still equivocal at best.

The inconsistent effects of valence on purchase decisions can be explained by factors related to consumer characteristics, brand or product characteristics, and textual content of reviews. Consumer characteristics can significantly moderate the relationship between star ratings and consumer purchases. According to Zou et al. (2011), the impact of valence is less salient for consumers with higher expertise. In addition, consumers rely on reviews whose valence are consistent with their initial beliefs about the product (Yin et al. 2016). Furthermore, novice buyers have higher purchase rate as well as higher return rate on extremely positive valence, compared to experienced buyers (Minnema et al. 2016).

Brand and product characteristics are also important moderators. Minnema et al. (2016) show that the impact of extreme positive reviews on return rate is higher for cheaper products than it is for more expensive products. Similarly, Zhu and Zhang (2010) find that review star ratings are more influential in consumer purchases when consumers are more experienced and the product is less popular. Finally, brand equity moderates the effect of star ratings on consumer purchases (Ho-Dac et al. 2013). In particular, positive (negative) reviews increase (decrease) the sales of weak brands but they have no significant impact on those of strong brands.

Regarding textual content, Schlosser (2011) reveals that consumers trust reviews in which valence and arguments are consistent with each other. In other words, they prefer one-sided review (i.e., only positive or negative statements are included) with an extreme star rating and two-sided review (e.g. both positive and negative statements are discussed) with a moderate star rating. Furthermore, Chen and Lurie (2013) suggest that the presence of temporal contiguity cues (i.e., when review writing closely follows consumption) such as “today” and “just got back” reduce consumers’ tendency to attribute reviews to the reviewer. This effect is stronger for positive reviews than negative reviews. Therefore, temporal contiguity cues in review texts attenuate the negativity bias. Similarly, Reimer and Ben (2016) also suggest that review content moderates the impact of star rating. Particularly, the authors show that consumers perceive review texts with argumentation (versus without argumentation) to be more trustworthy. More importantly, the relationship between star ratings and consumers’ purchase intention is stronger for trustworthy reviews than untrustworthy reviews.

These studies on review content seem to share one uniform view that consumers tend to rely less on those reviews that they attribute to the reviewers instead of product performances. In other words, if consumers believe that review content does not reflect true product performance but rather has been altered by reviewers due to personal factors such as their motivation, traits, and emotions, consumers will not trust the content of the reviews and thus will not use them to make their purchase decisions (Chen and Lurie 2013). However, would consumers always discount reviews based on those factors? Furthermore, are there any other factors beyond those identified in the extant literature that can cause consumers to discount reviews? These questions remain unanswered. The present paper tackle these questions by incorporating emotional arousal, another element of review content, as a new moderator of the relationship between star ratings and consumer decisions. Although recent research has discussed the role of emotional arousal (Yin et al. 2017, Ordenes et al. 2017), it has been treated as a stand-alone construct and its moderating effect on valence has not been investigated. Based on the emotion and the consumer decision making literature, the present paper suggests that emotional arousal can lead consumers to discount review valence only when consumers are at the search stage. In addition, when they are at their purchase stage, due to their anticipated regret, they are likely to engage in risk-averse behaviors. Therefore, even when negative reviews, especially the one-star ones, have high arousal, consumers will not discount them but rather will reject the product based on them. I demonstrate this differential moderating effect of emotional arousal through consumers' search and final purchase stages in three studies.

STUDY 1: EMOTIONAL AROUSAL EFFECT AT THE SEARCH STAGE

HYPOTHESIS DEVELOPMENT

A high review valence (e.g., five-star rating) indicates good product performance, which encourages consumer purchases. In contrast, a low valence (e.g., one-star rating) signals poor product performance and thus hinders product sales. Following this thought, a more extreme positive review (five-star ratings) can boost product purchase to a greater extent than a less extreme positive one (four-star ratings). Similarly, an extremely negative review (one-star ratings) can strongly discourage product purchases relative to a less extreme negative review (two-star ratings). Therefore, prior research has shown that the more extreme the valence of a review is, the more impactful the review is (Mudambi and Schuff 2010). For example, Pavlou and Dimoka (2006) report that extreme reviews of eBay sellers are more influential in consumers' decisions, compared with moderate reviews. Likewise, Forman et al. (2008) find that extreme reviews have greater impact on book sales than does less extreme reviews.

However, with respect to their purchase decision making process, consumers may also use emotional arousal to adjust this information reflected in star rating. Emotional arousal has been shown to influence people' thoughts and their behaviors (e.g., Bodenhausen et al. 1994) and product reviewers are not exceptions. Particularly, people with high arousal emotions (e.g., happy or angry) engage in less systematic information processing and have more biased judgments than those with low arousal emotions. For example, Bodenhausen et al. (1994) showed that angry subjects reported more stereotypic judgments than sad subjects, who experienced lower levels of emotional arousal. The scholars also observed that angry subjects relied on heuristic cues in a persuasion situation to a greater extent than sad subjects. Accordingly, it is likely that product reviewers assign a star rating for a product solely based on heuristics and their biased judgments that do not truly reflect product performance. In addition, prior research shows that when people can have a more complex cognitive perspective towards an object, they tend to give less extreme evaluation of the object (Linville & Jones 1980, Linville 1982). Meanwhile, high levels of emotional arousal can cause a reduction in cognitive complexity (Mano 1992). As a consequence, strongly aroused reviewers will assign an extreme rating for a product. These ratings are indicative of reviewers' lack of self-control rather than product performance (Parrott 1995). Following the above thought process, consumers may perceive high emotional arousal reviews as inaccurate evaluations of product performance and thus not informative.

When consumers have a planned purchase in their mind, they will engage in information search that later helps they make their purchase (Moe 2006). Online reviews are undoubtedly an important source for them to choose products for their consideration set (Hoffman and Novak 1996). The more informative the reviews are, the more likely consumers will make their decisions based on them. Search stage is defined as the stage under which consumers' main goal is to find acceptable alternatives to form a consideration set rather than to find the best choice (Shocker et al. 1991). To come up with a consideration set, they have to examine a large number of alternatives (Moe 2006). Consequently, efficiency and effort minimization are important goals at this stage, and consumers' attention tends to be highly selective (Bettman et al. 1998). Simpler decision strategies such as elimination-by-aspects and lexicographic strategies are more likely to be used. They are likely to use peripheral and salient attributes to make decisions. Review arousal that is usually accompanied by capitalized letters and exclamation marks is a salient attribute that consumers can utilize at this stage (Bradley et al. 1992). Utilizing review arousal to make decisions, consumers will find reviews with high arousal to be less informative due to reviewers' bias and lack of self-control. The discounted information in turn makes the impact of star rating less pronounced. In other words, the presence of high arousal emotions dilutes the negative star ratings' ability to discourage consumer purchases as well as the positive star ratings' ability to boost consumer purchases through consumers' perceived informativeness. Collectively, the discussion leads to the following set of hypotheses:

Hypothesis 1: *The effect of valence extremity on consumers' consideration set choices will be weaker when emotional arousal of the extreme review is higher than when it is lower*.

Hypothesis 2: *Consumers' perceived informativeness will mediate the effect of emotional arousal on valence impact. Specifically, high emotional arousal in a review will lead to lower perceived informativeness, which in turn leads to lower effect of review valence on consumers' consideration set choices*.

METHODOLOGY

Design

An experiment featuring with a 2 (rating valence: positive versus negative) x 2 (emotional arousal: high versus low) between-subjects design was conducted to test H1-H2. In each condition, consumers were asked to choose between two products, each accompanied by a review. One was the decoy review which was unemotional. The other review had a more extreme rating (e.g., more positive or negative) and was manipulated with emotional arousal. Naturally, a review with a more positive (negative) rating, everything else being equal, should be more likely to be chosen (rejected). However, I expected that high emotional arousal would weaken this valence effect.

Stimuli

A pretest was conducted to select appropriate reviews for this study and the following study. Since laptops have been used in previous emotion research (e.g., Kim and Gupta 2012), I retrieved ten moderately positive reviews for laptops from retail websites to develop stimuli. All emotional words and exclamation marks were removed from the reviews. Thirty Mechanical Turk employees in the United States were recruited to rate the arousal level and the informativeness of the reviews. Two positive reviews were selected as the decoy and the review of focus for my studies. They both were neutral in terms of emotional arousal ($M_{decoy} = 4.09$, $M_{main\ review} = 4.22$, $t = .54$, $p = .59$). They were also equally informative ($M_{decoy} = 5.52$, $M_{main\ review} = 5.5$, $t = .54$, $p = .59$). The arousal and the informativeness scales are in the measures section. The decoy review had a 4.0 rating. Negative adverbs were added to the review to create the negative version of the decoy with a 2.0 rating. The main review was manipulated with emotional arousal. It had 5.0 and 1.0 ratings for the positive and negative conditions, respectively. The stimuli are available in Appendix A. Respondents in the main study later were supposed to select one from two products. In the negative condition, one product had the manipulated one-star review and the other had the decoy two-star one. In the positive condition, one product had the manipulated five-star review and the other had the decoy four-star review.

To manipulate arousal, I added emotional words from the Affective Norms for English Words (ANEW) dictionary to the review (Bradley and Lang 1999). The dictionary has high reliability and validity and thus is the best known and the most frequently used dictionary in emotion research (e.g., Schmidtke et al. 2014, Eilola and Havelka 2010, Fujita et al. 2006, Kousta et al. 2011, Robinson and Tamir 2005). The dictionary originally has 1,034 words. Later, Warriner and colleagues modified it, creating a new dictionary with almost 14,000 words (Warriner et al. 2013). Each word has an arousal score ranging from 1 as the lowest to 9 as the highest. Positively emotional words (e.g., pleased) were added to the positive review and negatively emotional words (e.g., disappointed) were added to the negative review. Within each rating valence condition (i.e., positive and negative), emotional words had either low (e.g., pleased) or high arousal (e.g., awesome). In addition, exclamation marks and word capitalization were also included in the high arousal conditions (Yin, Bond, and Zhang 2016, Ordenes et al. 2017).

Participants and Procedures

One hundred and ten respondents from Amazon Mechanical Turk (Mturk) in the United States were recruited (M_{age} = 42.05, 82% Caucasian, 49% Female). They participated in my research in exchange for monetary compensation. I informed respondents that the main purpose of my study was to understand consumers' reactions to online reviews. Respondents were asked to imagine that they are searching for products for later laptop purchase. I then asked them to examine the reviews for two laptops with similar specifications to determine which one they will consider further for later purchase (Appendix B). Both reviews were either positive or negative. One laptop was the decoy that had the positive or negative unemotional review. The other laptop had the manipulated review with either low or high arousal. After examining the two laptops and their reviews, respondents indicated their purchase intention, perceived informativeness, and perceived review valence and arousal for each of the two products. They also answered demographic questions including age and education. Finally, they indicated how realistic the scenario was.

Specifically, respondents indicated the extent to which “The scenario is realistic” (1 = “strongly disagree”, 7 = “strongly agree”). Most of the respondents believed the scenario to be realistic (M = 5.96).

Measures

Perceived arousal: In order to check my manipulation of emotional arousal, I asked respondents “How do you think the reviewer was feeling at the time he/she writes the review?” (Yin et al. 2013). Respondents answered this question by rating three 7-point items including passive-active, mellow-fired up, and low-high energy (Berger 2011).

Perceived valence: Consumers should perceive the negative reviews to be negative and those positive reviews to be positive. For the manipulation of valence type, respondents indicated how the reviewer was feeling about the product at the time he/she writes the review with a seven-point item (-3 = negative, 3 =positive) (Berger 2011).

Informativeness: The three 9-point items from Gily et al. (1998) and Kim and Gupta (2012) were adopted. Those items were “The user review was useful”, “I think I learned a lot about the reviewed laptop after reading the user review”, and “The user review provided valuable information” (1 = strongly disagree, 7 = strongly agree). Whether a specific review was informative about product performance could vary among individual respondents. Therefore, reflecting this variable was the gap between the perceived informativeness of the manipulated review and that of the decoy review. The higher the number, the more informative the manipulated review was, compared with the decoy review.

Perceived distance along the purchase journey: to check the manipulation of the search stage, respondents indicated the extent to which they were close to their purchase. Specifically, they read “based on the scenario, when do you think the laptop purchase is likely to happen?” (1 = “in the next few days”, 2 = “within 2 weeks”, 3 = “within the next month”, 4 = “within the next 1-3 months”, and 5 = “longer than the next 3 months”).

Purchase intention: In the conditions with positive reviews, respondents were asked which one they were more likely to choose to purchase. In the conditions with negative reviews, respondents indicated the product they were more likely to reject for their purchase. This question was adapted to the different rating conditions in order to make the question more logically reasonable, since consumers would be more likely to reject products with negative reviews than to choose them. The question was on a semantic differential scale (0-100) with the two products as the two ending points. Specifically, the manipulated review with the more extreme valence was on the 100 end and the decoy review was on the 0 end. The higher the score, the more influence the valence had on consumers’ purchase decisions.

Pretest

A pretest was run to check the manipulation of arousal. One hundred and nine MTurk respondents in the US were recruited to evaluate the arousal of the reviews described in the stimuli section. An ANOVA was conducted with respondents’ perceived arousal as the dependent variable, and type valence (positive versus negative), arousal level (low versus high), and their interaction as the three independent variables. The effect of arousal level was significant ($F_{1,103}$ = 27.18, p <.001). Participants in the high arousal condition believe the manipulated reviews to be more aroused, compared with those in the low arousal condition ($M_{low\ arousal}$ = 5.15, $M_{high\_arousal}$ =6.31). Meanwhile, the coefficient of valence type was insignificant ($F_{1,103}$ = .28, p = .60); its interaction with arousal level was also not significant ($F_{1,103}$ = .36, p = .55). In addition, another

ANOVA was run with the same independent variables and perceived valence as the dependent variable. The effect of valence type was significant ($F_{1,103} = 1010.20$, $p < .001$). Reviews in the negative valence condition were perceived to be negative and those in the positive conditions were considered to be positive ($M_{negative} = -2.53$, $M_{positive} = 2.85$). Neither the main effect of arousal ($F_{1,103} = .002$, $p = .97$) nor its interaction with valence type ($F_{1,103} = .56$, $p = .46$) was significant.

Manipulation Checks

To check my manipulation of *emotional arousal* on the main sample, I ran an ANOVA with perceived arousal as the main dependent variable, and arousal level and valence type as the explanatory factors. The effect of arousal on perceived arousal was significant ($F_{1,105} = 23.69$, $p < .001$). Respondents who were assigned to the high arousal condition perceived the manipulated review to be more aroused, compared with those who were in the low arousal condition ($M_{low\ arousal} = 5.19$, $M_{high\ arousal} = 6.30$). The main effect of valence type was not significant ($F_{1,105} = .14$, $p = .71$). The interaction between arousal level and valence type was also insignificant ($F_{1,105} = .08$, $p = .79$). For valence, a similar ANOVA was performed with *perceived valence* as the dependent variable. The effect of valence type on perceived valence was significant ($F_{1,105} = 4149.90$, $p < .001$). Those who were assigned to the negative valence conditions believe the reviews to be negative and those who participated in the positive valence conditions perceive the reviews to be positive ($M_{negative} = -2.7$, $M_{positive} = 2.8$). The main effect of arousal was insignificant ($F_{1,105} = 1.75$, $p = .19$). The interaction between valence type and arousal level on perceived valence was also insignificant ($F_{1,105} = 3.69$, $p = .06$). Finally, the perceived stage of the respondents was also checked. Most of the respondents believed that the purchase would not likely occur soon ($M_{stage} = 3.56$).

RESULTS

In this study, the moderating effect of emotional arousal on consumers' reactions to valence extremity is captured through the extent to which people favor (reject) the product with the 5-star (1-star) rated emotional review over the product with the less extreme 4-star (2-star) rated and unemotional decoy review. If emotional arousal truly dilutes the star rating effect (H1), we would expect a weaker relative preference (rejection) for the 5-star (1-star) product over the 4-star (2-star) product when the 5-star (1-star) product review contains high arousal text than when it contains low arousal text. To test this hypothesis, I conducted an ANOVA with consumers' relative purchase/rejection intention likelihood for the more extremely rated product as the dependent variable, and valence type (positive vs. negative) and emotional arousal conditions as the two independent variables. The interaction between the two independent variables was not significant ($F_{1,\ 106}$ =1.5, p =.22). The main effect of valence was also insignificant ($F_{1,\ 106}$ =.30, p =.58). Meanwhile the effect of arousal was statistically significant ($F_{1,\ 106}$ =5.34, p <.05). Additional contrast analysis showed that the effect of valence extremity was lower when reviews had higher arousal than when they had lower arousal ($M_{low\ arousal}$ = 81.72, $M_{high\ arousal}$ = 71.05, p <.05). Figure 1 illustrates the effect of arousal on consumers' use of reviews with extreme valence.

Insert Figure 1 about here

H2 states that consumers' perceived informativeness of the review to the reviewer mediates the above effect of emotional arousal. In order to test this hypothesis, I first investigated whether high emotional arousal causes low perceived informativeness. An ANOVA was conducted with informativeness as the dependent variable, and valence type and emotional arousal conditions as the two independent variables. The interaction between the two independent variables was not significant ($F_{1, 106} = 3.31$, $p = .07$). The main effect of valence type was also insignificant ($F_{1, 106} = 1.44$, $p = .23$). Meanwhile, the effect of arousal on perceived informativeness was significant ($F_{1, 106} = 9.83$, $p < .01$). Further contrast analysis indicated that the extreme reviews with low arousal were more informative than the decoy reviews but the extreme reviews with high arousal were less informative than the decoy ones ($M_{low\ arousal} = .59$, $M_{high\ arousal} = -.08$, $p < .01$). Next, a regression of purchase/rejection intention on perceived informativeness was conducted. The model had a good fit (adjusted $R^2 = .11$). Informativeness had a significantly positive effect on consumer choices ($\beta_{infor} = 6.8$, $p < .001$). The more informative the reviews were, the more influential they were towards the choices of the respondents.

Then, as an overall test of my process model specified in H2, a moderated mediation analysis was performed through the bootstrapping method developed by Preacher and Hayes (2008). The bootstrap was set to 1,000 samples at the 95% confidence interval. This mediation analysis includes purchase/reject intention as the dependent variable, emotional arousal as the independent variable, valence type as the moderator, and informativeness as the mediator. Under both the positive and negative valence conditions, the indirect effect of emotional arousal on purchase intention through perceived informativeness was significant ($\beta = -4.27$, 95% CI = -8.77, -.92). The direct effect of emotional arousal was no longer significant ($\beta = -5.57$, 95% CI = -13.86, 2.53), suggesting the full mediation effect of informativeness. Therefore, the results supported H2

DISCUSSION

The results of this study show that emotional arousal has an important role in the impact of valence. Specifically, even though extremely negative reviews strongly discourage consumers' choices and extremely positive reviews boost consumers' choices, the effect is weaker when these reviews have high emotional arousal. This finding is consistent with the results reported by Berger and his colleagues (2010). Particularly, negative reviews are not always negative; rather, under certain situations, they are helpful for the firm's sales. The mediation analysis in this study also explains for this moderating factor. Reviews with higher emotional arousal tend to be less informative and consumers use informativeness to make decisions at the search stage. Therefore, they do not rely on extreme reviews with high arousal to choose products for their consideration set. Nevertheless, the effect of emotional arousal is likely to vary contingent on consumers' purchase stage because their thought processes tend be different between the search and the purchase stage. Thus, the purpose of the next study is to examine the moderating effect of emotional arousal in both stages.

STUDY 2: EMOTIONAL AROUSAL EFFECT AT DIFFERENT DECSIION STAGES

HYPOTHESIS DEVELOPMENT

The presence of high arousal emotions in a review may not always dilute its informational value and impact in many situations. An important consideration in this respect is where the reader is along the purchase path. Consumers generally have two common goals when making their decisions. They want to maximize the accuracy of their decisions and minimize their cognitive effort for those decisions (Bettman et al. 1998). Yet, one goal comes at the expense of the other. In other words, in order to obtain high accuracy, consumers have to spend a great amount of cognitive effort. Meanwhile, if consumers want to minimize their cognitive effort, they have to sacrifice accuracy. When consumers are at their search stage, they are willing to sacrifice accuracy for cognitive effort due to the large number of products they have to examine. Because high arousal emotions are highly salient and memorable (Bradley et al. 1992), they are likely to serve as an easy heuristic for eliminating undesirable alternatives or keeping potentially desirable candidates. Thus, as shown in the first study, high arousal leads reviews to be less informative about product performance and thus less impactful in consumers' consideration set choices during the early stage.

Consumers' decision strategy is different when consumers are at the purchase stage. Purchase stage is defined as the stage under which consumers choose the best choice for their purchase from the consideration set gathered from the search stage (Sambandam and Lord 1995). Since consumers make a final choice at this stage, anticipated regret emerges, and more compensatory and complex decision rules are likely to take over in evaluating the consideration set (Simonson 1992, Betty et al. 1998). Even though high arousal leads consumers to perceive less informativeness from reviews, they do not always make choices based solely on review informativeness. Rather, to avoid future regret of selecting wrong products, they extend further cognitive effort to think about the reviews beyond these reviews' informativeness. For positive reviews, selecting a product with a four-star review rather than another one with a high arousal five star review is not likely to cause much anticipated regret since the former is still a good one even when the latter is in fact better. Therefore, consumers still reject the high arousal positive reviews due to its low level of informativeness.

However, review informativeness is not impactful under the negative review situation. Specifically, although a high arousal one-star review is less informative about product performance than a two-star review, keeping the product with the one-star review will cause much regret if that product turns out to be truly bad. Furthermore, being risk-averse at this stage, consumers want to avoid negative consequences from buying the product and high arousal negative feelings are one among them. They try to avoid post-purchase dissonance and do not want to get mad or frustrated with their bought items and (Schmidt and Spreng 1996). Therefore, unlike their behavior at the search stage, consumers at the purchase stage are more likely to reject the one-star reviews if those reviews have high arousal.

Hypothesis 3: *Emotional arousal has a differential effect along the purchase journey. During the search stage, emotional arousal weakens the effect of review valence on consumers' consideration set choices. During the purchase stage, emotional arousal a) weakens the effect of positive valence on consumers' purchase choices but b) strengthens the effect of negative valence on consumers' purchase choices.*

METHODOLOGY

Design and Procedures

To test H3, I conducted an experiment with a 2 (valence type: positive versus negative) x 2 (emotional arousal: low versus high) x 2 (purchase stages: search stage versus final purchase stage). Two hundred and fifty Mturk respondents in the United States were recruited to participate in this study. Similar to study 1, they were informed that the study was to understand consumers' reactions to online reviews. Respondents were exposed to the reviews for two laptops; one was the decoy and the other had its review manipulated with emotional arousal. In the purchase stage, respondents were asked to choose one of the two products as the final choice for their purchase. Meanwhile, in the search stage, they were asked to select one of the two products for further consideration later. The complete scenarios are available in Appendix B. The procedure were similar to study 1.

After their exposure to the two products, respondents indicated their purchase intention, their perceived informativeness, and their perceived review valence and arousal for each of the two products. These measures were the same as in Study 1. With respect to purchase intention in the search stage, respondents indicated which of the two products they were more likely to choose (reject) to consider further for later purchase. In the purchase stage, respondents indicated which one they were more likely to choose (reject) for their purchase. Finally, respondents answered the scenario realism item and demographic questions.

Manipulation Checks

A t-test was conducted to analyze the effect of purchase manipulation on perceived stage. Purchase manipulation had a significant effect ($t=8.23$, $p <.001$). Those who were in the search condition expected to make a purchase later than those were in the search stage ($M_{search} =3.11$, $M_{purchase} =1.80$). The manipulation of arousal and valence were also checked to see whether whether the purchase/search stages affected consumers' perception of review valence and arousal. Thus an ANOVA was performed with perceived arousal as the dependent variable, and arousal level, valence type, purchase stage, and their interactions as the independent variable. Arousal level has a significant effect ($F_{1,240} = 42.12$, $p < .001$). Participants in the high arousal condition perceived the manipulated review to be more aroused ($M_{low\ arousal} = 5.23$, $M_{high\ arousal} = 6.21$). Meanwhile, the effect of valence ($F_{1,240} = 1.33$, $p=.25$) and purchase stage ($F_{1,240} = 1.82$, $p=.18$) and their interactions were statistically insignificant. Similarly, another ANOVA was conducted with perceived valence as the dependent variable. The effect of valence was strongly significant ($F_{1,240} = 1629.98$, $p < .001$). Arousal ($F_{1,240} = 1.09$, $p=.30$) as well as purchase stage ($F_{1,240} = 3.47$, $p=.06$) did not have any significant impact on perceived valence.

RESULTS

To test H3 and retest H1 and H2, I conducted an ANOVA with consumers' purchase/reject intention as the dependent variable, and valence type, arousal level, purchase stage, and their interactions as the independent variables. The three-way interaction was insignificant ($F_{1,242} = 2.28$, $p < .05$). There was a significant interaction between purchase stage and arousal level ($F_{1,242} = 5.28$, $p < .05$). Further planned contrast analyses were conducted for both search and purchase stages. The effect of arousal in the search stage was significantly negative ($M_{low\ arousal}$ = 78.70, $M_{high\ arousal}$ = 61.59, p <.001). There was no significant difference between the two levels of arousal in the purchase stage ($M_{low\ arousal}$ = 73.06, $M_{high\ arousal}$ = 72.78, p=.96). However, when an ANOVA was performed with purchase/reject intention as the dependent variable, and arousal level, valence type, and their interaction as the independent variables, there was a significant interaction between arousal level and valence type ($F_{1,123}$ = 12.15, p < .001). Further planned contrast analyses showed that arousal weakens the effect of extremely positive reviews ($M_{low\ arousal}$ = 83.91, $M_{high\ arousal}$ = 67.29, p <.05), yet it strengthens the effect of extremely negative reviews ($M_{low\ arousal}$ = 60.71, $M_{high\ arousal}$ = 79.03, p <.05). The results thus supported H3. Figure 2 illustrates the reported findings.

Insert Figure 2 about here

To test the mediation effect of arousal level on consumer choice at the search stage, I first conducted a t-test to examine the relationship between arousal level and informativeness. The effect was significant (t = 3.85, p < .001); high arousal leads consumers to perceive extreme reviews less informative than decoy reviews ($M_{low\ arousal}$ = .67, $M_{high\ arousal}$ = -.39). A regression conducted also showed that informativeness increases the impact of extreme reviews ($\beta_{informativeness}$ = 8.79, p < .001, adjusted R^2 = .19). A bootstrap mediation analysis with 1,000 samples and a 95% confidence interval was then run. Consistent with study 1, the indirect effect of emotional arousal through informativeness was significant (β= -8.06, 95% CI = -13.51, -3.49) and the direct effect of emotional arousal was no longer significant (β= -6.64, 95% CI = -24.69, 11.18) confirming the full mediation effect of informativeness during the search stage.

Since there was an interaction between valence and arousal on consumer choice at the purchase stage, the effects of both arousal and valence on informativeness were examined. An ANOVA was first conducted with informativeness as the dependent variable, valence type, arousal level, and their interaction as the independent variables. The interaction term between valence type and arousal level was not significant ($F_{1,123} = 2.42$, $p=.12$). The main effect of valence type was also insignificant ($F_{1,123} = 1.43$, $p=.23$). Meanwhile, the main effect of arousal level was significant ($F_{1,123} = 10.09$, $p < .01$). Regardless of whether the reviews were negative or positive, those with high arousal were perceived to be less informative ($M_{\text{low arousal}} = .62$, $M_{\text{high arousal}} = -.19$, $p <.01$).

To test the effect of informativeness on consumer choice in the purchase stage, I ran a regression with consumer purchase/reject intention as the dependent variable, and valence type and informativeness as the explanatory factors. The model had a good fit (adjusted $R^2 = .14$). There was a significant interaction between valence type and informativeness ($\beta_{Va*Infor} = 10.38$, $p <.01$). When valence was positive, informativeness about the product increased the impact of extreme reviews ($\beta_{Va} = 10.24$, $p <.01$). Such an effect was insignificant in the negative review condition ($\beta_{Infor} = -.14$, $p =.96$). Separate bootstrap mediation analyses for the negative and the positive conditions were conducted. Each of the bootstrap was set to 1000 samples with a 95% confidence interval. As expected, for the positive reviews, the indirect effect of arousal through informativeness was significant ($\beta = -11.28$, 95% CI = -19.20, -4.29). The direct effect of arousal was insignificant ($\beta = -4.97$, 95% CI = -13.10, 3.36), indicating full mediation effect of informativeness. Meanwhile, for the negative reviews, the indirect effect of arousal through informativeness was not significant ($\beta = -.05$, 95% CI = -4.25, 4.45). Thus, informativeness about product performance mediated the effect of arousal among extremely positive reviews, but it did not affect consumers' use of extremely negative reviews.

DISCUSSION

This study confirms that the moderating effect of emotional arousal on review valence is different along consumer purchase journey. At the search stage, consumers simply use heuristics to make choices for their consideration set. Emotional arousal in review text that is reflected by strong emotional words and exclamation marks tends to be salient. Therefore, it becomes a heuristic for consumers. In particular, consumers perceive extreme reviews with higher emotional arousal less informative. Since informativeness about product performance is an important factor directing them on how to choose products, consumers discount extreme reviews with high arousal. Meanwhile, during the purchase stage, consumers engage in a more complex decision-making process. For positive reviews, informativeness about product performance still plays an important role. Meanwhile, for negative reviews, anticipated regret leads people to make decisions independent of informativeness. In other words, even though consumers believe that extremely negative reviews with higher arousal are less informative, they still reject the products associated with these reviews because of their anticipated regret of ordering these products. Although the study showed insightful findings, it occurred in a lab experiment setting. Therefore, a field study (study 3) that closely examines consumers' actual online behaviors through the data of a large retailer can potentially improve the robustness of the current findings.

STUDY 3: EMOTIONAL AROUSAL EFFECT TESTED IN A FIELD STUDY

METHODOLOGY

Whereas the two experimental studies provided new insights about emotional and valence, they may not completely reflect what consumers would do in their actual purchase journey. In addition, whereas star ratings can reflect review valence, such a reflection is not fine grained since consumers cannot choose any point between any two adjacent ratings. Therefore, to measure the effect of valence, I estimated review valence from review texts in this study. Emotional valence and star rating was strongly correlated with each other ($\rho = .215$, $p < .001$). The present research attempted to replicate the earlier findings by examining real-world behaviors recorded on a major UK retailer website. The data contained information pertaining 35,206 products under two broad categories – Technology (e.g., tablets, printers, phones, MP3 players) and Home and Garden (e.g., pillowcases, mattresses, mirrors). It included the browsing actions (e.g., pages viewed and reviews exposed to) as well transactions made by 243,000 consumers during a two-month period (February and March 2015). Together these consumers generated approximately 2.5 million total page views, 12.3 million review impressions, and 30,000 purchases in the aforementioned product categories.

Two important steps were followed before the main analysis of emotional arousal effect. First, for emotional arousal values for the reviews, the reviews' texts were coded following previous studies on sentiment analysis. Review valence was also measured using the same procedure. Second, since the effect of emotional arousal vary depending on consumers' purchase stages as stated in my hypotheses, these stages were identified. The following sections discuss the detailed description of these two steps.

Emotional Arousal and Valence Coding

To code the level of emotional arousal and valence reflected in each review, I first completed text preprocessing tasks including tokenization, part-of-speech tagging, and lemmatization (e.g., Feldman et al. 1998). Tokenization involves dividing review sentences into different linguistic units such as words, punctuations, numbers, and alpha-numerics. Next, part-of-speech tagging automatically classified the words in the reviews based on morpho-syntactic categories such as verbs, nouns, and adjective. For instance, “distractions” should be classified as a noun, yet “distracts” should be considered a verb. Then these words were lemmatized through the removing of their inflectional forms. For example, the word “distractions” and “distracts” were converted back to “distraction” and “distract”, respectively. This process allowed later matching between the review text and the emotion dictionary without encountering error messages. The ANEW dictionary described in Study 1 was used to compute emotional arousal scores. Prior research suggests that the dictionary works for not only American speakers but also British speakers (Eilola and Havelka 2010). It was thus suitable for my UK-based data. To calculate emotional arousal scores of individual reviews, I matched emotion words in the reviews with those in the dictionary and extracted the mean arousal scores for that word from the dictionary. For review valence, since the dictionary has valence scores as well, review text was also matched with the valence scores from the dictionary to capture valence value for each review.

It should be noted that a weighted mean of the arousal and valence scores of all emotional words within a review was used to calculate the emotional arousal level of that review (Ramaswamy 2011). This was due to the fact that certain words showed greater variance in ratings across individuals than other words in the study originally used to develop the ANEWS dictionary. Accordingly, the arithmetic mean did not reflect the emotion scores well. For instance, “angry” had an arousal score of 7.17 out of 9 and the standard deviation of this score was only 2.07 indicating that most human coders for the dictionary development agreed that “angry” was a high arousal emotion. Meanwhile, “depressed” had an arousal score of 4.72 and its standard deviation was 2.95 which implied that there was less agreement among the coders regarding the arousal level of this emotion.

To adjust for the above issue, I used the probability density function of a normal distribution to measure the probability of the word’s score rating falling exactly at the mean. I then used those probability as weights when summing mean scores. As an example, if a review had two emotional words – angry and depressed, employing the arithmetic mean approach would yield an arousal score of [(7.17+4.72)/2] = 5.945. However, by utilizing the probability density function, I observed that the probability that 7.17 was the true mean of “angry” was .193 and the probability that 4.72 was the true mean of “depressed” was .135. Based on the total of these two probabilities, the weight of 7.17 in the review was (.193/.328) = .588 and that of 4.72 was (.135/.328) = .412. So, the total arousal score of the review was (7.17*.588+4.72*.412) = 6.16. This total score leaned more towards “angry” than towards “depressed”. Thus, this calculation helped me get arousal and valence scores that better reflect reviewers’ feelings than those from the arithmetic mean method.

Purchase Stage Identification Process

To identify each consumer’s decision-making stage at the time of exposure, I drew from previous studies of clickstream data that showed distinct browsing patterns for different decision contexts (Moe 2003). In particular, previous research has shown that consumers who are in the deliberate search stage will examine many products within a limited number of categories, and will not spend too much time on each page due to the fact that they are still early in the purchase process. In contrast, those in the final purchase stage will have narrowed down candidate products to a much smaller consideration set. Hence they view a few products likely from a single category. The measures used in Moe (2003) were adopted to characterize each browsing session, as shown in Table 4 of the paper. With these variables, k-means cluster analysis was performed to classify the browsing sessions into different clusters. The analysis started with two clusters, and added more clusters into the solution until the added cluster became very small or until the added cluster was very similar to existing clusters (Moe 2003). Before the cluster analysis, outlier sessions which had extremely long duration or/and super high page views were removed. Based on the results of the cluster analysis, browsing sessions that fall within the pattern of search stage and purchase stage were retained for the main analysis.

RESULTS

Purchase Stage Classification

The store visits were cluster-analyzed using nine different criteria including 1) the number of pages viewed, 2) the duration of session visit, 3) the percentage of category pages over the total number of pages viewed, 4) the percentage of unique product pages over all product pages viewed, 5) the percentage of unique category pages over all category pages viewed, 6) the ratio of product pages over category pages, 7) the number of times a product page was viewed within the same session, 8) the percentage of features consumers use per page (e.g., viewing questions and answers related to a product), 9) the percentage of features consumers interacted with (e.g., sorting reviews based on overall review volume). The amount of time spent per page was similar among the clusters and thus it was not used to divide the clusters. The mean statistics were summarized in

Table 1. The results indicated three main clusters that had between subject sum of squares account for 82% of total sum of squares. The three clusters were different in a variety of ways. For people who are at the late purchase stage, since they were ready to buy products, they tent to spend more time on product pages than on different product category pages (64.75% over total pages viewed). Compared to the other two clusters, they spent more time during their visit (36.68 minutes). Among the three clusters, they also had the highest times of repeating the same product pages during a session (3.71 times). They were also more likely to use page features than the other clusters (64.39% of available features). This suggests that these customers focused on completing their purchase task on the retailer's website. Therefore, their conversion rate was also the highest (5%), among the three clusters.

Insert Table 1 about here

For people who were at the early purchase stage, they also showed their purchase effort through their browsing behavior. They spent on average 13.37 minutes per session. The number of times they repeat the same product pages per session was also high (3.07 times). They utilized features available on pages. Specifically, they typically used 52.36% of the available page features. Their conversion rate was high but smaller than that of the late stage buyers (4%). Meanwhile, for people who belonged to the search stage, as they expected to check other stores and probably came back later, they were less committed to their purchase task on the retailer's site. On average, they spent only 1.39 minutes per session. In addition, each visit had only 3.73 pages. The times they repeat the same product pages were also limited (1.6 times per session). They used only 35.09% of page features. Therefore, their conversion rate was the lowest among the three clusters (1.5%).

Main Analysis

For each cluster, a logistic regression was run to test the hypotheses. The model is represented as follows:

$$\Pr(Purchase_{ijk}) = \frac{\exp(u_{ijk})}{1+\exp(u_{ijk})} \quad (1)$$

$$u_{ijk} = \beta_0 + \beta_1 VA_{i+} + \beta_2 VA_{i-} + \beta_3 EA_i + \beta_4 SR_{i+} * EA_i + \beta_6 SR_{i-} * EA_i + \beta_7 SR_{i+} * PS_i + \beta_8 SR_{i-} * PS_i + \beta_9 EA_i * PS_i + \beta_{10} SR_{i+} * EA_i * PS_i + \beta_{11} SR_{i-} * EA_i * PS_i + \sum_l \alpha_l X_{ijkl} + \sum_m \gamma_m X_{ijm} + \sum_m \theta_n Z_{jm} \quad (2)$$

The dependent variable was the eventual purchase outcome of product *j* due to exposure to review i in session k. Specifically, as shown in equation 1 and 2, the purchase outcome was a function of the utility u_{ijk} which was determined by review *i*'s positive valence (VA_+), negative valence (VA_-), emotional arousal (EA), and their interactions. The value of positive valence was any positive valence score of reviews which were between zero and five. The value of negative valence was any negative valence score of reviews that fell between zero and negative five. It should be noted that separate variables for positive and negative valence were included to take into account the asymmetric effect of emotional arousal suggested in the hypotheses. In addition, this approach was consistent with previous research (e.g., Chevalier and Mayzlin 2006). Besides these focal variables, a set of covariates that could also affect the purchase decision were also included. These included session-specific review controls (X_{ijkl}) such as the number of other reviews on the same screen as the focal review and position of the focal review in the list of reviews on the same screen, and session-constant review controls (X_{ijm}) such as length and readability of the review. Some product controls (Z_{jm}) such as price, description length, average review rating and/or percentage of reviewers recommending the product, and total review volume (total number of reviews) were also added to the equation (e.g., Mudambi and Schuff 2010; Wu 2013). The summary of the results are presented in Table 2.

Insert Table 2 about here

For the early state purchase, the model had a good fit (McFadden R^2 = .295). The main effect of arousal was negatively significant ($\beta_{arousal}$ = -.046, p < .05). The interaction effect of positive valence was positive and significant ($\beta_{pos\_valence*\ arousal}$ = .025, p < .05), which suggested that the negative impact of arousal was still negative (β = -.021) but smaller if review valence was higher. In other words, for less positive reviews, if they had emotional arousal, they could be discounted even more, compared to those with higher valence. In sum, the results confirmed that consumers were less likely to buy products associated with positive reviews when reviews had high emotional arousal. Meanwhile, the interaction effect of negative valence and arousal was insignificant ($\beta_{neg\_valence*\ arousal}$ = .012, p = .46). The findings therefore suggested that consumers were less likely to buy products associated with high arousal and negative reviews, regardless of how negative they were. To conclude, consumers discounted high arousal positive reviews but utilized high arousal negative reviews. The results therefore were consistent with my findings for the purchase stage. Whereas the second study examined the effect of emotional arousal at the purchase stage with extreme reviews only and kept the moderate reviews as unemotional decoy ones, the findings of this study generalize the effect of arousal to moderate reviews when those reviews had high emotional arousal.

For late stage buyers, the model also had a good fit (McFadden R^2 = .295). However, the main effect of arousal was negative yet insignificant ($\beta_{arousal}$ = -.016, p = .66). Its interaction with negative valence was also insignificant ($\beta_{neg\_valence*\ arousal}$ = .010, p = .51). However, the interaction between arousal and positive valence was significantly positive ($\beta_{pos\_valence*\ arousal}$ = .058, $p < .05$). The main effect of valence was negative and insignificant ($\beta_{pos\_valence}$ = -.02, p = .97). The positive total effect (β = .56) suggested that high arousal helps positive reviews boost product purchases. Since these late stage buyers were committed to buying products at the retailer's site within their visit and thus probably made a decision in their mind by then, they looked for affirmative evidence to their already-made choices (Chevalier and Mayzlin 2006). Therefore, high arousal from those positive reviews plays an important role in supporting their confirmation bias.

The regression model in the search stage also had a good fit (McFadden R^2 = .307). In this model, the main effect of arousal was insignificant ($\beta_{arousal}$ = -.011, p = .56). Its interaction terms with positive valence ($\beta_{pos\_valence}$ = -.003, p = .78) and with negative valence ($\beta_{neg\_valence*\ arousal}$ = -.010, p = .43) were also not significant either. Yet, the effect of negative valence was significantly negative on customers' decisions to purchase at this stage ($\beta_{neg\_valence}$ = -.044, $p < .05$). In other words, the more negative the reviews were, the less likely consumers would be buying the products. The results in this model were inconsistent with what I found in the first two studies. However, it should be noted that one of the limitations of the data was that it did not capture consumers' consideration set. The use of purchase decision as the dependent variable for the search stage then limited the ability of the study to generalize the effect of emotional in this particular stage.

DISCUSSION

The effects of arousal in the purchase stage found in this study are consistent with the results in Study 2. In particular, consumers discount positive reviews with high emotional arousal. Furthermore, the negative effect of arousal is even more severe for less positive reviews. Meanwhile, consumers rely on negative reviews with high arousal and the effect is consistent across different levels of negative valence. It is also interesting to observe the effect of emotional arousal when consumers are committed to making a purchase and already make a decision in their mind regarding what product to buy. Specifically, due to consumers' confirmation bias at this particular point, high arousal improves the ability of positive reviews in encouraging product purchases. Yet, the effect of emotional arousal was minimal at the search stage if consumers' purchase decision was used as the main dependent variable.

GENERAL DISCUSSION

Through two experiments and a field study, the paper show that emotional arousal plays different role depending on which stage the customers belong to. Specifically, when consumers are at their search stage and do not expect to make their purchase soon, extreme reviews with high emotional arousal are less impactful in consumers' selection of products for their consideration sets. Meanwhile, when consumers are at their purchase stage and expect to make a purchase soon, the effect of emotional arousal is different. They discount positive reviews with high emotional arousal, and the effect is smaller for more extreme reviews. Yet, for negative reviews, due to their anticipated regret, consumers are likely to follow arousal to make decisions and are inclined to reject products associated with high arousal reviews. The first two studies examine such effects for extreme reviews only since it is more reasonable to see high arousal associated with extreme reviews than with moderate reviews. The results of the field study then generalize the effects to less extreme reviews in the purchase stage. Interestingly, when consumers are ready to make a purchase and make a decision in mind, emotional arousal enhances the effect of positive reviews on consumers' purchase decision due to their confirmation bias.

THEORETICAL IMPLICATIONS

The present paper provides several important contributions to the online word-of-mouth literature. First of all, it helps to reconcile the inconsistent effects of valence documented in prior research (King et al. 2014). If reviews are extremely negative or positive, they should have more effect on consumers' decisions than moderate reviews. However, some previous studies find only the effect of extremely negative reviews (e.g., Cui et al. 2012), whereas others lean towards the impact of extremely positive reviews (e.g., Chevalier and Mayzlin 2006). Such equivocal conclusions can be due to the high emotional arousal level usually associated with these negative reviews. Particularly, if the consumers are at their search stage and they are not planning to purchase any product soon, they will discount extreme reviews with high emotional arousal. However, when they plan to make a purchase soon, they discount positive reviews with high arousal yet base their decision on negative reviews with high emotional arousal. For moderately positive reviews which are already less influential than extremely positive reviews, if they have high emotional arousal, their effect was even more limited. This also confirms the finding documented by Schlosser (2011). Specifically, Schlosser (2011) shows that reviews whose content and star ratings are inconsistent with each other are not valuable to consumers. Likewise, moderately positive or negative reviews with high emotional arousal belong to that type of reviews. Finally, if consumers are ready to buy a product and form a decision in their mind already, positive reviews with high arousal turn out to be helpful in confirming consumer's choice. This is consistent with the confirmation bias reported earlier by Chevalier and Mayzlin (2006). In short, emotional arousal helps explain the equivocal effect of review valence on consumer decisions found in prior research.

In addition, previous research on online word-of-mouth has focused mainly on consumers' purchase decision such as product sales (e.g., Chevalier and Mayzlin 2006, Duan et al. 2008); yet, with a few exceptions (e.g., Li et al. 2011), it did not give much attention to the other stages of consumer purchase journey. As shown in the current paper, consumers' thought process can be starkly different depending on which stage they are at. Specifically, when consumers are at their search stage, due to the high number of products and brands they have to examine, consumers are likely to use heuristics and thus utilize salient features of reviews such as emotional arousal

(Bettman et al. 1998). Meanwhile, when consumers are planning to make a purchase soon, they are more risk-averse. Their anticipated regret will prevent them from selecting products based purely on heuristics, especially if those products are associated with negatively high arousal reviews. Further, when consumers are at the late purchase stage identified in the field study, consumers already form a decision in their mind and are ready to order a product, they examine reviews not to compare products but just to confirm their choices. Therefore, examining different stages of consumers' decision-making process provides a more complete picture of how impactful online reviews are.

Furthermore, the paper also contributes to the text mining literature in the Information Technology field. Sentiment analysis in this field primarily focuses on emotional valence (e.g., Hu et al. 2012). Specially, sentiment analysis allows researchers to identify whether a particular review sentence is positive or negative about the product or the firm. However, consumers' experiences with product use can be reflected by not only emotional valence but also emotional arousal. Emotional arousal thus gives a more complete picture of the important role of emotions in online reviews. Therefore, the extraction and analysis of emotional arousal from this paper helps researchers better understand what consumers think and feel through review text. Emotional arousal can also be examined under other contexts such as social media posts, which deserves further research attention.

Finally, the paper confirms the importance of emotional arousal as a source of information for consumers' decisions stated in extant literature (Mayer 2001). In particular, according to the emotions-as-information theory, for example, if we feel pleasant when we are around somebody, then such an emotion leads us to like that person (Wyer and Calston 1979). As another instance, Esch et al. (2012) conjecture that consumers use emotions to evaluate brands. Similarly, Li et al.

(2011) also report that consumers used their emotions formed from their impression of a retailer's website to determine whether they exchange their personal information with the retailer. However, these studies examine the effect of emotional arousal experienced within the focal consumers. The present paper suggests that consumers indeed interpret the level of emotional arousal of other people. They then use it to make inferences about the quality of the information provided by these people. Hence, not only the emotional feeling of the decision makers but also the emotions of the information providers can be informative for the decision makers.

MANAGERIAL IMPLICATIONS

As indicated by the findings of the three studies, not all extremely positive reviews are beneficial and similarly, not all extremely negative reviews are damaging. Instead, when consumers are at their search stage, they are likely to discount these extreme reviews if they are highly aroused. Therefore, positive reviews exposed to consumers at this stage should have less arousal; meanwhile, firms should not be worried about extremely negative reviews with high arousal. On their websites, firms usually tend to present only positive reviews to customers, which sometimes lose their credibility in consumers' eyes. In fact, when firms track the browsing of their consumers and they can identify that some consumers are first time visitors and are likely to be at their search stage, firms should present both positive and negative reviews. This will give consumers the impression that firms are objective when presenting these reviews without harming the firm's ability to sell. In addition, social media such as Facebook are those channels through which consumers interact with the firm to gather information at the initial stage and make their purchases later on other channels. Consumers are less likely to make purchases directly on those channels (Kapko 2016). Therefore, positive reviews on Facebook and other social media platforms should be less aroused to gain trust from these customers.

Meanwhile, based on consumers' web browsing behavior, if some consumers appear to be repeat visitors of the firm's website, the firm should avoid presenting extremely negative reviews with high arousal to them since these reviews will sway consumers away from the firm and its products. Yet, if consumers are ready to make a purchase, especially those who abandon the shopping cart and then come back, firms should expose these consumers to extremely positive reviews with high arousal either through reminder emails or through the list of product reviews on the firm's website.

LIMITATIONS AND FUTURE RESEARCH

Although the paper employs both lab experiments and a field study to test the role of emotional arousal on the relationships between review valence and consumer choices, it encounters several limitations that can be directions for future research. First, whereas the two lab experiments allow me to test the moderating effect of arousal in both the search and purchase stages, the data in the field study does not provide much information about the consideration set of consumers at their search stage. Therefore, I find corroborated evidence from the field study regarding the impact of emotional arousal at the purchase stage, yet cannot find supportive results for the search stage. Future research thus should delve into the search stage if data are comprehensive enough to examine the role of arousal on consumers' consideration sets.

Second, whereas the field study provides very interesting information regarding consumers' behaviors at their late purchase stage, the first two lab experiments only look at the search and the early purchase stage. Particularly, I did not ask respondents to assume that they already chose one of the products in their mind and then examine the reviews of these products to determine whether they would change their decision or not. As a result, future research should also observe this particular stage in a lab setting experiment to confirm the findings of my field study.

Third, in the first two lab experiments, time was a factor used to manipulate the search and the purchase stages. It is due to the fact that people who are at the search stage extend their effort to acquiring product information as long as the cost of doing so is less than the benefits potentially obtained from such information (Greenleaf and Lehmann 1995). Thus, there is usually a time delay between the search stage and the purchase stage (Greenleaf and Lehmann 1995). However, in certain situations, consumers engage in the purchase stage immediately after their search stage. Future research should examine the effect of emotional arousal under those situations.

Finally, the paper does not capture the effect of personal differences. The way that consumers interpret emotions can vary among individuals. For example, consumers' regulatory focus can play a role here. At the late purchase stage where consumers tend to have confirmation bias, such as bias can be greater among promotion-focused consumers than it is among prevention-focused consumers since the former has a stronger motivation to make a purchase than the latter (Higgins 1998). In addition, personality can influence how consumers interpret emotions of reviewers and make according inferences about the provided information. As an instance, neuroticism refers to a person's emotional instability, which are represented by insecurity, anxiousness, and hostility (Barrick and Mount 1991). People who are high in neuroticism are thus likely to discount reviews at the search stage to a greater extent than those who has more emotional instability. In addition, due to their anxiousness, the extent to which they look for affirmative evidence for a product choice made in their mind is even stronger, compared with other people. Therefore, individual differences warrant future research attention.

REFERENCES

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology, 44*(1), 1-26.

Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science, 22*(7), 891-893.

Bettman, J. R., & Park, C. W. (1980). Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *Journal of Consumer Research*, *7*(3), 234-248.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*(3), 187-217.

Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology*, *24*(1), 45-62.

Bonnet, D., & Nandan, P. (2011). Transform to the Power of Digital: Digital Transformation as a Driver of Corporate Performance. *Report, Capgemini Consulting*.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (pp. 1-45). Technical report C-1, the center for research in psychophysiology, University of Florida.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 379-390.

Celsi, R. L., & Olson, J. C. (1988). The role of involvement in attention and comprehension processes. *Journal of Consumer Research*, *15*(2), 210-224.

Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research*, *50*(4), 463-476.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345-354.

Feldman, R., Fresko, M., Kinar, Y., Lindell, Y., Liphstat, O., Rajman, M., ... & Zamir, O. (1998). Text mining at the term level. *Principles of Data Mining and Knowledge Discovery*, 65-73.

Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology*, *90*(3), 351.

Gilly, M. C., Graham, J. L., Wolfinbarger, M. F., & Yale, L. J. (1998). A dyadic study of interpersonal information search. *Journal of the Academy of Marketing Science*, *26*(2), 83-100.

Greenleaf, E. A., & Lehmann, D. R. (1995). Reasons for substantial delay in consumer decision making. *Journal of Consumer Research*, *22*(2), 186-199.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In *Advances in Experimental Social Psychology* (Vol. 30, pp. 1-46). Academic Press.

Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, *52*(3), 674-684.

Kim, J., & Gupta, P. (2012). Emotional expressions in online user reviews: How they influence consumers' product evaluations. *Journal of Business Research*, *65*(7), 985-992.

King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing*, *28*(3), 167-183.

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General*, *140*(1), 14.

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, *19*(4), 456-474.

Linville, P. W. (1982). The complexity–extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology*, *42*(2), 193.

Linville, P. W., & Jones, E. E. (1980). Polarized appraisals of out-group members. *Journal of Personality and Social Psychology*, *38*(5), 689.

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, *77*(1), 87-103.

MacKenzie, S. B., & Lutz, R. J. (1989). An empirical examination of the structural antecedents of attitude toward the ad in an advertising pretesting context. *The Journal of Marketing*, 48-65.

Minnema, A., Bijmolt, T. H., Gensler, S., & Wiesel, T. (2016). To Keep or Not to Keep: Effects of Online Customer Reviews on Product Returns. *Journal of Retailing*, *92*(3), 253-267.

Moe, W. W. (2006). An empirical two-stage choice model with varying decision rules applied to internet clickstream data. *Journal of Marketing Research*, *43*(4), 680-692.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com.

Niedenthal, P. M. (2008). Emotion concepts. *Handbook of Emotions*, *3*, 587-600.

Parrott, W. G. (1995). But emotions are sometimes irrational. *Psychological Inquiry*, *6*(3), 230-232.

Pechmann, C., & Stewart, D. W. (1990). The effects of comparative advertising on attention, memory, and purchase intentions. *Journal of Consumer Research*, *17*(2), 180-191.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123-205.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*(2), 135-146.

Robinson, M. D., & Tamir, M. (2005). Neuroticism as mental noise: a relation between neuroticism and reaction time standard deviations. *Journal of Personality and Social Psychology*, *89*(1), 107.

Russell, J. A., & Pratt, G. (1980). A description of the affective quality attributed to environments. *Journal of Personality and Social Psychology*, *38*(2), 311.

Sambandam, R., & Lord, K. R. (1995). Switching behavior in automobile markets: a consideration-sets model. *Journal of the Academy of Marketing Science*, *23*(1), 57-65.

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

Schmidt, J. B., & Spreng, R. A. (1996). A proposed model of external consumer information search. *Journal of the academy of Marketing Science*, *24*(3), 246-256.

Shocker, A. D., Ben-Akiva, M., Boccara, B., & Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing Letters*, *2*(3), 181-197.

Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research*, *19*(1), 105-118.

Unnava, H. R., Agarwal, S., & Haugtvedt, C. P. (1996). Interactive effects of presentation modality and message-generated imagery on recall of advertising information. *Journal of Consumer Research*, *23*(1), 81-88.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.

Wu, Philip Fei (2013), "In search of negativity bias: An empirical study of perceived helpfulness of online reviews," *Psychology & Marketing*, 30 (11), 971-84.

Yin, D., Bond, S. D., & Zhang, H. (2016). Keep Your Cool or Let It Out: Nonlinear Effects of Expressed Arousal on Perceptions of Consumer Reviews. *Journal of Marketing Research*.

Yin, D., Bond, S., & Zhang, H. (2013). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly, 38(2),* 539-660.

Zou, P., Yu, B., & Hao, Y. (2011). Does the valence of online consumer reviews matter for consumer decision making? The moderating role of consumer expertise. *Journal of Computers*, *6*(3), 484-488.

TABLES AND FIGURES

TABLE 1: STUDY 3 - CLUSTER ANALYSIS

Cluster	Search Stage	Early Buying Stage	Late Buying Stage
N	245821 (81.9%)	43967 (14.64%)	10451 (3.48%)
Pages	3.73	10.02	11.80
Session Duration (in second)	83.66	801.94	2200.97
% of Category Pages over All Pages	43.63	40.42	34.83
% of Unique Product Pages over All Product Pages	34.94	45.43	47.48
% of Unique Category Pages over All Category Pages	58.42	78.60	76.60
Ratio of Product Pages over Category Pages	1.23	1.71	1.76
# of Times A Product Page is Viewed	1.59	3.07	3.71
% of Features Used Per Page	35.09	52.36	64.39
% of Feature Interactions Per Page	58.42	78.60	76.60
Purchase Likelihood	.015	.04	.05

TABLE 2: STUDY 3 - LOGISTIC REGRESSION RESULTS

DV: Purchase decision	Shallow	Early	Late
Positive Valence	0.000	-0.003	-0.024
	(0.016)	(0.020)	(0.036)
Arousal	-0.011	-0.046*	-0.016
	(0.019)	(0.023)	(0.036)
Negative Valence	-0.044*	-0.005	0.009
	(0.022)	(0.024)	(0.054)
Negative Valence * Arousal	-0.010	0.012	0.010
	(0.013)	(0.015)	(0.024)
Positive Valence * Arousal	-0.003	0.025*	0.058*
	(0.010)	(0.013)	(0.024)
Review Length	0.000	-0.002	-0.002
	(0.001)	(0.001)	(0.001)
Product Price	-0.006***	-0.005***	-0.002
	(0.001)	(0.001)	(0.001)
Review Readability	0.001	-0.001	0.002
	(0.001)	(0.001)	(0.001)
Number of other reviews in the same session	0.009	-0.001	0.027*
	(0.011)	(0.010)	(0.013)
Review location	0.002	0.003	-0.001
	(0.002)	(0.002)	(0.003)
Product Average Rating	0.192**	0.252**	0.311
	(0.064)	(0.083)	(0.164)
Product Review Volume	0.000***	0.000**	0.000
	(0.000)	(0.000)	(0.000)
ProductDescription Length	0.000	0.000	-0.001
	(0.001)	(0.001)	(0.002)
Constant	-4.621***	-4.372***	-4.775***
	(0.294)	(0.363)	(0.714)
N	245,821	43,967	10,451
McFadden	0.307	0.295	0.295

Notes: Standard errors in parentheses; *** significant at 0.001, ** significant at 0.01, * significant at 0.05

FIGURE 1: STUDY 1 - EMOTIONAL AROUSAL EFFECT AT THE SEARCH STAGE

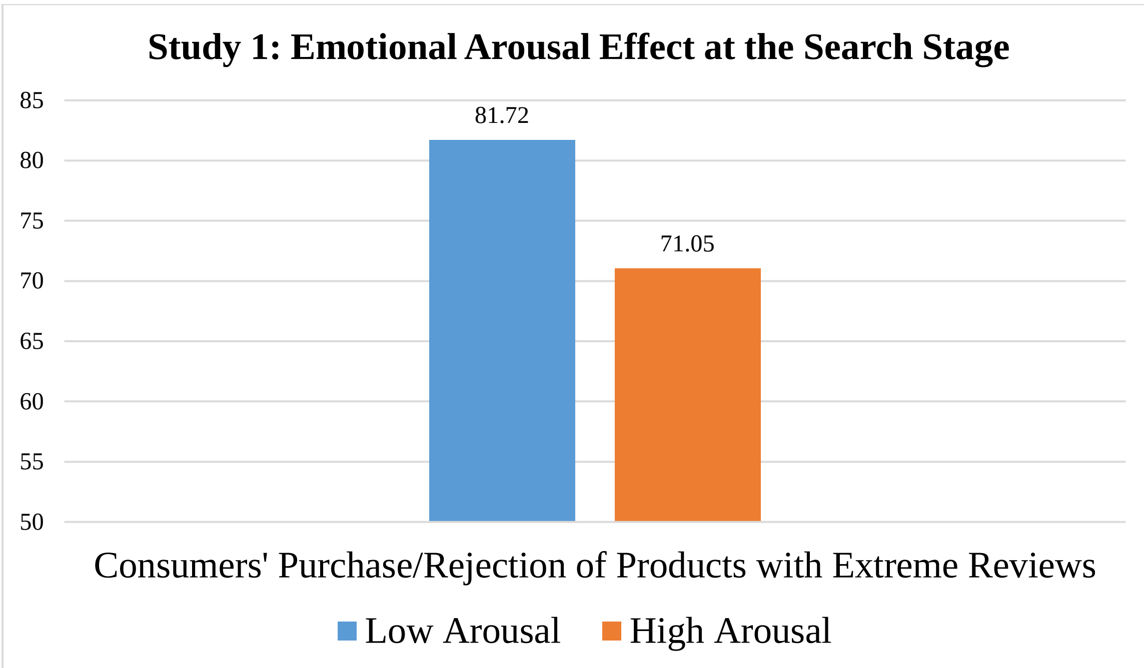
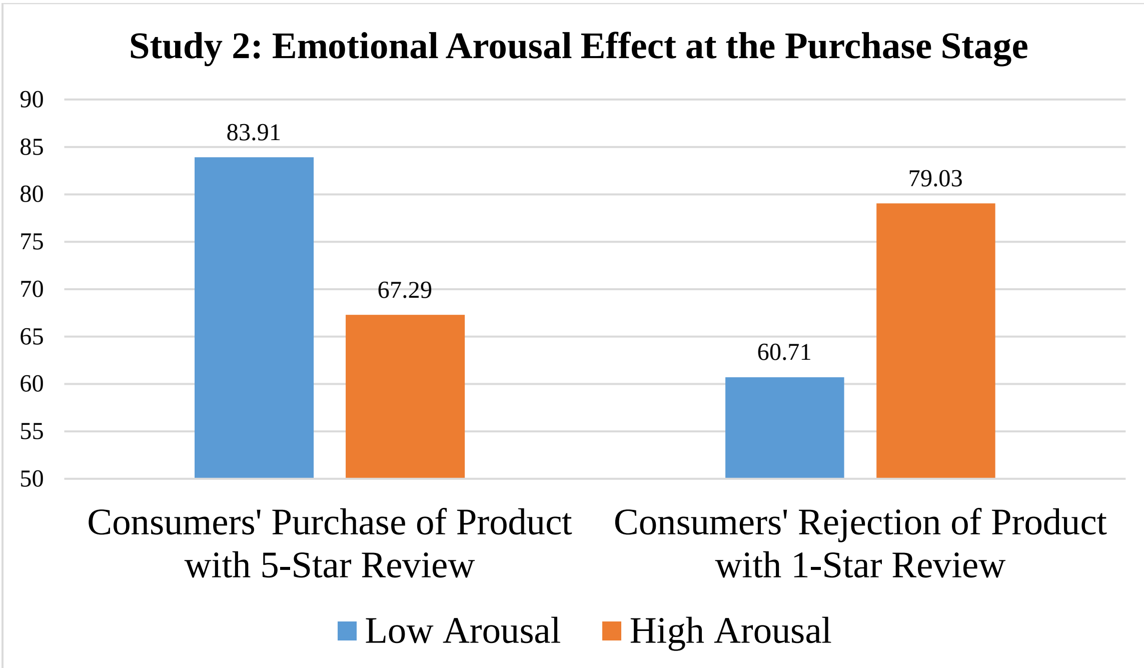
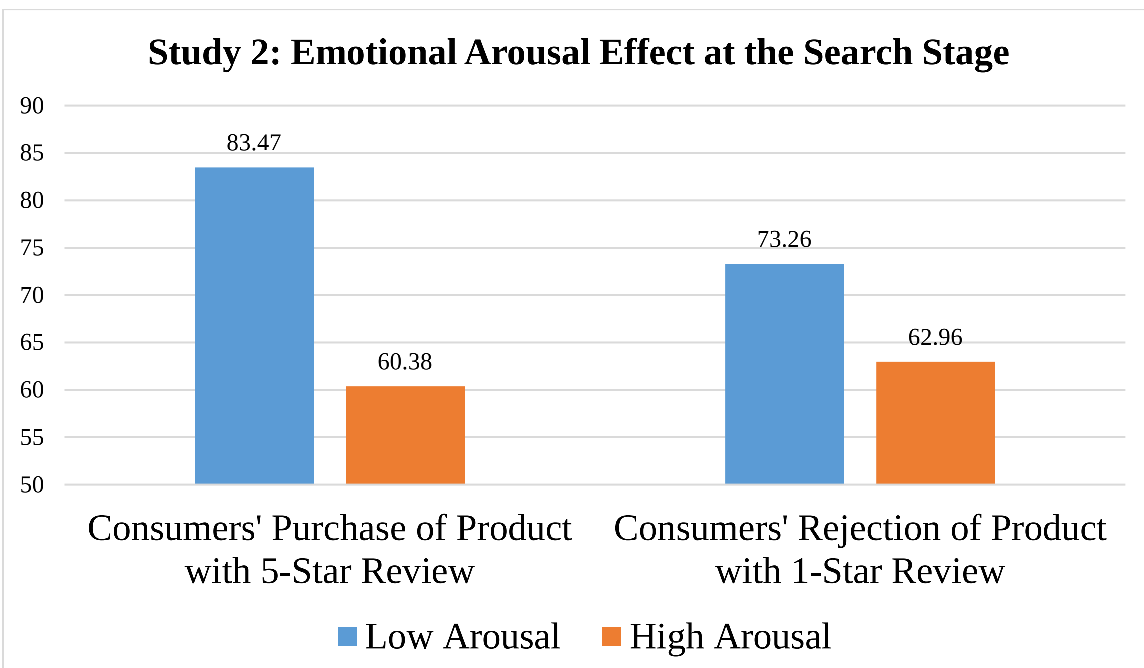


FIGURE 2: STUDY 2 - EMOTIONAL AROUSAL EFFECT AT DIFFERENT DECISION STAGES





APPENDICES

APPENDIX A – REVIEW STIMULI

	Negative Valence	Positive Valence
Decoy Review (2 or 4 Star Rating)	This is a poor laptop. The construction seems to be cheap. It allows the screen to lie flat to 180 degrees, if you desire, but it is not a touch screen. The display resolution is difficult to read and the machine runs loudly due to the poorly made hard drive and has a slow response time.	This is a good laptop. The construction seems to be sturdy. It allows the screen to lie flat to 180 degrees, if you desire (it is not a touch screen though). The display resolution is clear and easily readable and the machine runs quietly due to the solid state hard drive and has a quick response time.
Manipulated Review - Low Arousal (1 or 5 Star Rating)	Got this laptop for work and home and so far is not working well. Slow processor and poor screen. Construction quality is low, keyboard is awkward at best, speakers are poor, and the battery life sits around only 3-4 hours or real world usage. The mouse pad is not of centered so accidentally hit several times. I'm so disappointed with this purchase.	Got this laptop for work and home and so far is working perfectly. Fast processor and clear screen. Construction quality is excellent, backlit keyboard is nice, speakers are very nice, and the battery life sits around 8-9 hours or real world usage. The mouse pad is of centered so not accidentally hit. I'm so pleased with this purchase.
Manipulated Review - High Arousal (1 or 5 Star Rating)	AWEFUL computer!!! Got it for work and home and so far is not working well. Slow processor and poor screen. Construction quality is extremely low. I HATE the keyboard, it is awkward at best, speakers are poor too, and the battery life sits around ONLY 3-4 hours or real world usage. This is terrible!!! The mouse pad is NOT of centered so accidentally hit several times. I'm so UPSET with this purchase!!!	AWESOME computer!!! Got it for work and home and so far is working perfectly. Fast processor and clear screen. Construction quality is excellent. I LOVE the backlit keyboard. The speakers are very nice too, and the battery life sits around 8-9 hours or real world usage. This is so cool!!! The mouse pad is of centered so not accidentally hit. I'm so THRILLED with this purchase!!!

APPENDIX B – PURCHASE STAGE SCENARIOS

Search Stage:

"Imagine that you need a new laptop within the next year. You have just started researching laptops to consider for your later purchase. You would like to narrow down the possibilities to a few products for further research. Please examine the reviews for the two laptops below and answer questions that follow. Please note that these two laptops have similar specifications."

Purchase Stage:

"Imagine that you need a new laptop within a week. You have researched laptops for a while and already formed a list of a few laptops that meet your requirements. You expect to make a final purchase today. Please examine the reviews for the two laptops below and answer questions that follow. Please note that these two laptops have similar specifications."

ESSAY 2

VALENCE FRAMED IN PERCENTAGES – WHEN THE RICH GETS RICHER

ABSTRACT

Given the heavy influence of online review volume on consumers' purchases (Liu 2006), products newly introduced to an online store are often at a great disadvantage compared with competing incumbents. Even if the newly introduced product has a somewhat higher quality than an existing product, many consumers will still choose the existing product for the reliability and popularity signaled by its high review volume. It is therefore of strategic interest for the makers of newly introduced products to overcome the disadvantage and motivate consumers to consider their products more equally on their merit. In this research, through five lab experiments, I show that this can be achieved by changing the way one frames a product's recommendation rate. In particular, the dominant effect of a high review volume can be attenuated by presenting recommendation rate of products as numbers (e.g., 44 out of 50 consumers recommend the product) instead of as percentages (e.g., 88% of 50 consumers recommend the product).

INTRODUCTION

Online reviews are gaining importance in determining consumers' purchase decisions since many consumers trust them as much as personal word-of-mouth (DeMers 2015). Nevertheless, possessing positive reviews alone may not be enough for a product to get the attention of consumers. Rather, due to the prevalence of online reviews as well as the potential manipulation of reviews made by sellers, consumers also consider how many reviews a product has accumulated. The more reviews a product has, the more reliable the valence of the product reviews should be (Zhu and Zhang 2010). Following this preference, Amazon now even allows consumers to sort products not only by average valence but also by the number of reviews through the installation of the Amazon™ Sort App. Consumers' reliance on review volume has also been documented extensively in the online review literature (e.g., King et al. 2014, Liu 2006). Unfortunately, consumers' preference for high review volume creates a significant disadvantage for newly introduced products that may have higher quality but fewer online reviews than their established competitors. How can these products overcome this disadvantage?

Extant research has thoroughly examined review valence, volume, and their relative importance in consumers' decision making process. Some studies argue that review volume affects product sales whereas review valence does not (Liu 2006, Duan et al. 2008). Meanwhile, other studies suggest that the effects of review volume and valence are contingent on other factors such as product types, review sites (You et al. 2015), firm characteristics (Bla and Sturman 2014), and whether national data or market-level data are used to analyze the effects (Chintagunta et al. 2010). Although these studies provide insightful findings regarding the role of volume and valence, they do not provide specific strategic advice on how to help reduce the weight of review volume and boost that of review valence to increase the sales of new products when consumers prefer review volume to review valence. In other words, how new products can tackle their disadvantage in terms of low review volume remains unanswered.

My research therefore aims to solve the above problem by examining the effect of valence framing on consumer purchase decisions through five lab experiments. Valence, which reflects positive or negative product evaluation, can be expressed not only in a 5-star or 10-star system but also as binary choices such as recommend or not (e.g., Gupta and Harris 2010) and thumbs up or down. In practice, some companies such as Youtube, Netflix, and Uber have decided to move from a rating scale system to thumbs up/down with the hope that the new system is less confusing and prone to bias (O'Donovan 2017). The valence measure in such a binary choice setting then is the extent to which people recommend or thumbs-up. Unlike 5-star ratings that have more or less the same format across platforms, binary choice-based valence can be in either percentage (e.g., 60% of 50 customers recommend this product as used by Lowes) or absolute numbers (e.g., 30 out of 50 customers recommend this product as used by Yankee Candle). Drawing from the number framing literature, I posit that the presence of these two different formats in review valence paves the way for new products to overcome its review volume disadvantage.

The findings from this research show that consumers engage in two distinct approaches when processing the two numerical formats mentioned above. If review valence is in the percentage format, consumers adopt a piecemeal approach which allows them to compare two products based on valence and volume separately. The difference in review valence between the two products becomes overshadowed by the difference in review volume. Consequently, the product with a higher review volume and yet lower valence is likely to be chosen. In contrast, when review valence is in the absolute number format, consumers adopted a holistic approach in which they examine both review valence and volume within each product before comparing between products. This approach reveals that the product with a higher volume has more people who thumbed down or did not recommend. As a result, consumers' likelihood of choosing the product with a greater valence yet lower volume is greater.

The findings in this research provide three important contributions to marketing research and practice. First, previous studies on numerical framing usually examine consumers' purchase decisions for an individual product rather than allowing them to make comparisons between choices. The assumption under this holistic approach is that consumers form an overall impression about a product before comparing it with another product regardless of the representation format. This leads to the conclusion that consumers interpret portion information (e.g., 25% shipping surcharge versus $15.52, 70% fat versus 7.5g fat) based on the base number (e.g., original price $50, 10g fat in total) and tend to contrast the magnitude of the portion information to this base number without considering the format difference between the portion and the base numbers. My studies show that the opposite can be the case. By allowing consumers to choose between two products, the current findings suggest that consumers in fact utilize a piecemeal approach when the portion information is in a percentage format. These findings contribute to the numerosity literature and suggest the need to consider choice task format in examining number framing effects.

Second, by examining numerical framings of valence, my findings help to reconcile the effects of valence and volume discussed in the literature. In particular, the current research confirms previous findings regarding the dominant role of volume. Yet, my research also shows that *how* consumers process valence and volume can determine the effects of these two factors on consumers' purchases. Third, the findings suggest that firms should not randomly choose the format of review valence. Specifically, the percentage format causes products with higher review volumes to dominate over their competitors with lower review volumes. Meanwhile, the absolute number format helps attenuate the dominance of volume and boost the salience of valence. Therefore, firms should consider using the absolute number format if they want to provide opportunities for new products to reach consumers faster.

CONCEPTUAL BACKGROUND

THE RELATIVE IMPACT OF VOLUME AND VALENCE ON CONSUMER CHOICES

Due to abundant information as well as the availability of many brands in the market, consumers are likely to use heuristics to make decisions. Hence, aggregate review factors such as review volume (i.e., the total number of reviews) and valence (i.e., the tone or preference of reviews which is usually expressed as positive, negative, or neutral) can be helpful for consumers. Accordingly, review volume as well as its relative impact (compared to valence) on consumer purchase decisions have received a great deal of research effort (King et al. 2014).

Even though the impact of review volume is still debated (King et al. 2014), many studies agree that volume plays a crucial role in driving product sales. For instance, Yang et al. (2010) show that review volume has a positive impact on box office revenue and the effect is stronger for a mass product than a niche product. Similarly, investigating the impact of online reviews on TV show viewership, Cadario (2015) reports that review volume is not influential on the early episodes but its impact increases over time and then declines at the end of a show's life.

More interestingly, previous studies have also examined the impact of volume in relation to that of valence. For example, Liu (2006) find that review volume significantly determines box office revenue whereas valence does not have much influence. Likewise, Duan et al. (2008) suggest that valence has little persuasive impact on consumer choices, but volume positively affects box office sales. Showing more nuance in the relative impact of valence versus volume, Chintagunta et al. (2010) suggest that if aggregated national data are used, the effect of volume is strong but valence does not matter; however, if local market-level data are used, the finding is reversed. Meanwhile, Blal and Sturman (2014) show that the relative impact of volume and valence depends on the firm's characteristics. Specifically, review valence has a greater impact on luxury hotels, whereas review volume has a stronger effect on lower-tier hotels. Furthermore, a meta-analysis conducted by You et al. (2015) shows that both valence and volume are powerful determinants of consumer purchases and their effects are contingent on product types and review sites. Specifically, the effects of these two characteristics are stronger for privately consumed, low-trialability products offered by less competitive industries. Additionally, the impact of volume is higher for durable goods and for reviews on specialized review sites, and the impact of valence is more pronounced for community-based sites.

Whereas previous studies have extensively explored the impact of review volume on consumer purchase and have compared its effect with that of valence, those studies drew widely different conclusions and they focused mostly on contextual factors such as product characteristics and review sites. Little is known about how consumers process valence and volume differently due to the way the information itself is framed, even though framing alone has been shown to significantly influence consumers' perceptions and decisions (Chen et al. 1998, DelVeccio et al. 2007). When consumer ratings are done in a binary choice fashion (thumbs up/down or recommend/not recommend), valence can be expressed in either a percentage (e.g., 80% of 150 consumers recommend) or an absolute number format (e.g., 120 out of 150 customers recommend). Based on the numerical framing literature, I argue that how valence is framed can influence the relative salience of valence and volume and thus the impact of these two review factors on consumer purchase decisions.

NUMBER FRAMING AND ITS IMPACT ON CONSUMER DECISIONS

Choice of format (percentages versus absolute numbers) has been shown to have differential influence on consumer behavior in prior research (Chen et al. 1998, DelVeccio et al. 2007). Specifically, previous studies suggest that percentages are more difficult to evaluate than absolute numbers, leading consumers to use simplifying heuristics for the former. For example, Morwitz et al. (1998) posit that when surcharges for shipping are presented in percentages (versus absolute numbers), the resulting multiplication and addition operations based on the original price require more cognitive effort than only addition operations needed for absolute numbers. Accordingly, consumers tend to use low-effort heuristics or ignoring strategies to process the prices. As a result, they recall lower total costs when the surcharges are framed in percentages. Similarly, DelVeccio et al. (2007) suggest that consumers are more reluctant to calculate a revised price if the discount is in percentage term (versus as an absolute number).

Exploring the specific heuristics that consumers follow, Weathers et al. (2012) argue that consumers pay attention to raw magnitude and ignore the scale of the information. Therefore, they perceive, for example, the shipping charge of 28.5% to be larger than a shipping charge of $15.52 for a product priced at $54.47, even though the actual charge is exactly the same. As another example, Tangari et al. (2014) draw similar conclusions regarding consumers' perceptions of fat-content in food packaging. Specifically, they suggest that consumers with low numeracy ability will have a more favorable attitude toward a decrease in fat content when it is presented in a percentage format (i.e., 70% less saturated fat) than when it is in an absolute number format (7.5g less saturated fat) due to consumers' focus on unit magnitude and ignorance of scale. Furthermore, Tao and colleagues (2017) find that consumers consider a rating of 80 over 100 worse than a rating of 8 over 10 since consumers estimate the portion information against its closest anchor and the former rating shows a relatively larger absolute distance without reference to the scale difference. In summary, prior studies suggest that consumers process a piece of information differently when it is in percentage versus absolute number format.

Although there is extensive research on the framing of numbers, most of previous studies tend to focus on consumers' holistic approach in processing a single product's information. In other words, regardless of the framing structure (percentages or absolute numbers), consumers are assumed to evaluate the portion information (e.g., discount, surcharges, donations) based on an information anchor for the same product (e.g., base price) (e.g., Kleber et al. 2016, Morwitz et al. 1998, Weathers et al. 2012). Following this assumption, when faced with different products, consumers tend to process the portion information and the base anchor within individual products first, come up with their separate conclusion on each product's value, and then make their choice among available products based on that value. For instance, Morwitz et al. (1998) state that consumers consider a product's base price and its surcharges to come up with its total cost to determine their demand for that product. Weathers et al. (2012) also suggest that consumers consider the magnitude of a product's surcharge based on the magnitude of the product's base price.

One important limitation of this focus on a holistic approach using a single product in previous studies is their inability to observe when and how consumers may contrast product choices differently depending on the information representation format. In this study, I posit that consumers do not always utilize a holistic approach. Instead, sometimes consumers apply a piecemeal approach, where they compare products based on their portion information and on their anchor information separately (Sujan 1985, Muthukrishnan et al. 2001). Particularly, if the portion information and the anchor information represent two distinct product attributes, consumers are likely to use the piecemeal approach by comparing products based on each attribute (Sujan 1985). Such a piecemeal approach can be especially relevant in the online review context. For any product, retailers provide information about review volume and valence (i.e., recommendation and/or star ratings). Valence can be considered portion information in a binary rating system since it shows what percentage (or how many people) of the total review volume has positive ratings for the product (e.g., 90% or 180 out of 200 customers recommend the product). Prior research has shown that consumers may compare products based on valence and volume individually (e.g., Liu 2006, Lee et al. 2015), hinting at the possibility of a piecemeal approach. My research suggests that whether consumers use a holistic or a piecemeal approach can be determined by the way the firm frames review valence in a binary rating setting.

HYPOTHESIS DEVELOPMENT

According to the concreteness principle, individuals generally accept information in the format provided unless they are required to do otherwise (Slovic 1972, Weathers et al. 2012). Therefore, although 44 out of 50 is mathematically the same as 88% of 50, consumers will engage in two distinct processes evaluating the two information representations. The key difference between the two representations is whether review valence and volume are expressed in the same format (all numbers) or in different formats (percentages and numbers). I posit that consumers will use a piecemeal approach when valence and volume are expressed in different formats, while a holistic approach will be used when valence and volume are both expressed as absolute numbers.

PIECEMEAL APPROACH

According to Sujan (1985), when evaluating products based on information from sources such as advertisements or package labels, consumers examine each piece of information separately. This piecemeal approach is based on consumers' perceptions that products generally comprise of discrete attributes and that each attribute has a distinct subjective value (Anderson 1972, Fiske 1982). Regarding number framing, when portion information is expressed as a percentage and the total amount is presented as an absolute number on a quantitative scale, these two pieces of information will appear as two distinct attributes to consumers since their subjective values are deemed different. For instance, a valence score of 80% recommendation rate indicates the positivity of product performance whereas a volume of 200 total reviews signify product popularity with higher numbers indicating higher popularity (Duan et al. 2008, Liu 2006, Lee et al. 2015) and higher reliability of the reviews (Zhu and Zhang 2010). Unlike valence which has a ceiling value of 100%, volume can increase to any positive integer. Thus, valence and volume become two distinct product attributes due to the difference in their subjective values. Accordingly, when faced with two products for which valence is expressed as percentages rather than absolute numbers, consumers are likely to use a piecemeal approach by comparing the two products' valence and volume separately.

Following the above approach, if a product is recommended by 88% of 50 consumers and another product is recommended by 80% of 200 consumers, the low volume product will have a valence advantage (e.g., 8%) yet a volume disadvantage (e.g., 150 reviews). Similar to any two attributes of products (e.g., price and months of warranty), the difference in valence is not directly comparable to that in volume. Under such a difficult tradeoff, Pelham et al. (1994) suggest that consumers are tempted to employ a numerosity heuristic by examining the magnitude without regard to the size of the units. This is analogous to considering, for instance, a house of eight small rooms as larger than a house with five spacious rooms (Pelham et al. 1994). With respect to the product review example above, the use of the numerosity heuristic results in the perception that the difference in valence (8%) is small relative to the difference in review volume (150). This accentuates the advantage of the high-volume product.

HOLISTIC APPROACH

In contrast, when recommendation rates are presented in absolute numbers, all numbers (whether volume or valence) exist in the same format. Due to proximity and the general realization that the number of recommendations is relative to the number of total people who have expressed their opinions, consumers are likely to first compare the two numbers within each product (44 vs. 50) and then between the two alternative products. According to DelVecchio et al. (2007), consumers are likely to choose simple heuristics such as subtraction calculations rather than engaging in complicated exercises such as divisions. Supporting this view, Tao et al. (2017) show that consumers perceive an 8 out of 10 product rating to be better than an 80 out of 100 product rating because the former has a 2-unit gap between the rating and its anchor whereas the latter has a 20-unit gap. Following this logic, when comparing review valence and volume that are both expressed as absolute numbers, consumers are likely to use a subtraction exercise instead of a division. The tendency to use a subtraction exercise to find the gap is also consistent with consumers' negativity bias documented in previous studies (Baumeister et al. 2001). Specifically, when buying a product, consumers are likely to pay particular attention to reviews with negative valence. The gap calculated from the subtraction provides a concrete answer regarding the number of people who do not recommend a product.

A few consequences result from this within-product subtraction exercise. One, the volume information becomes absorbed into a within product comparison with valence, which reduces the obvious disadvantage of the low-volume product. Second, the subtraction will lead to the conclusion that fewer absolute individuals do not recommend the low-volume product (50-44=6) compared with the high-volume product (200-176=24). As the subtraction concerns potentially negative outcomes of purchase decisions, people also tend to be more risk seeking, favoring losses that are uncertain to sure losses (Kahneman & Tversky 1979). With high volume often considered as a signal of reliability (i.e., less risk), a low volume is preferred in the assessment of potentially negative outcomes. Taken together, the within-product subtraction exercise consumers are likely to engage in under the absolute number format can alleviate the disadvantage of a low-volume product and work in favor of such products. This leads to the following hypothesis:

Hypothesis 1: The positive impact of volume on purchase likelihood will be lower when valence is presented in absolute numbers than when it is in percentages.

OUTLINE OF STUDIES

There were totally five lab experiments conducted to examine the impact of review numerical framing. The aim of study 1a was to test H1 and demonstrate that an absolute number format (versus a percentage format) weakens the effect of review volume on consumers' purchase likelihood and such effect is consistent across different levels of volume gap between two competing products. Study 1b was performed to corroborate the effect observed in Study 1a by setting the valence level (i.e., recommendation rate) of the two products to be the same. In addition, consumers was exposed to the products sequentially rather than simultaneously. Study 2 provided a more stringent test of the effect by setting the volume of the low volume product option to be very small relative to that of the high volume choice. Furthermore, three different product categories were examined to generalize the effect of numerical format. An eye-tracking experiment was conducted in Study 3 to show the process under which consumers go through to process review valence and volume. Finally, in Study 4, H2 was tested to examine whether the color format of review valence and volume breaks the effect of absolute number framing.

STUDY 1A: NUMERICAL FRAMING AND PURCHASE LIKELIHOOD

METHODOLOGY

Design

My first study aimed to demonstrate that the positive impact of volume on purchase likelihood will be weaker when valence is presented in absolute numbers than when it is in percentages (H1). In the study, participants were asked to choose one product from a product pair characterized by valence and volume tradeoffs. The study has a 2 (valence format: percentage versus absolute numbers) x 2 (volume difference: small versus large) between-subjects design. The varied levels of volume difference between the two products was to ensure the robustness of the findings. Specifically, I expect to observe the ability of the absolute number format to reduce consumers' likelihood of choosing the high-volume choice even when the volume difference between the two products is large.

Pretest

To choose an appropriate level of valence for my stimuli, I conducted a pretest asking twenty-eight undergraduate students the minimum level of recommendation rate that a product should have in order to be considered for their purchases. The respondents were given different levels as possible answers (at 10% increments from 20% to 90% plus an "other" option). Many respondents (N=13, 46%) chose 80% or higher whereas each other level was selected by less than 17% of the sample. Therefore, I used 80% for my main study. This is consistent with He and Bond

(2015), which suggests that consumers are more likely to consider highly-rated products than low-rated products.

Participants and Procedures

A hundred and fifty-four Amazon Mechanical Turk (MTurk) respondents participated in the main study (mean age = 39.5, 51.3% female). The respondents were told that they would be making a hypothetical choice between two products based on the information provided. Shoes served as the main products in the study based on previous research (Punij 2012). Each participant was randomly assigned to one of the four conditions. In all conditions, information about two pairs of shoes was shown side by side. The displayed information included the recommendation rate and the review volume of each option. Figure 1 illustrates the display format. Recommendation rates were expressed as either percentage or raw numbers of consumers who recommended the product out of all consumers who rated it (i.e., review volume).

Insert Figure 1 about here

In the small volume difference condition, respondents had to choose between an option with 50 reviews and an 88% recommendation rate (44 customers recommended), and another option with 200 reviews and an 80% recommendation rate (160 customers recommended). In the large volume difference condition, the valence level remained the same but the high-volume product's total number of reviews was even larger at 450 total consumers.

After examining the information for the shoes, respondents were asked to select the option that they would be more likely to purchase. To check the manipulation of numerical framing, I asked respondents to indicate whether they agreed that the recommendation rates were presented in absolute numbers rather than in percentages (1 = “strongly disagree”, 7 = “strongly agree”). I also measured participants’ demographic information including their age, gender, and education. Finally, I asked respondents to indicate the extent to which they believed the scenario was realistic (1 = “strongly disagree”, 7 = “strongly agree”).

RESULTS

An ANOVA with perceived format as the dependent variable was conducted to test my manipulation of numerical format. The manipulation works as expected ($F_{1,146} = 66.18$, $p < .001$) Further planned contrast analysis showed that those in the absolute number format condition were more likely to agree that the valence was presented in absolute numbers than were participants in the percentage format condition ($M_{number} = 5.65$ and $M_{percentage} = 2.88$). For tradeoff type, its main effect ($F_{1,146} = .85$ and $p = .36$) and its interaction with numerical format ($F_{1,146} = .01$ and $p = .94$) were insignificant. Furthermore, respondents rated the realism of the scenario as high ($M = 5.69$), and the realism rating did not differ significantly across the four conditions ($F_{1,147} = 1.7$ and $p = .19$).

My main hypothesis predicts that consumers are less likely to choose a higher-volume product when the valence is in absolute numbers than when it is in percentages. I conducted a logistic regression with the binary choice variable (0 = low-volume choice, 1= high-volume choice) as the dependent variable and valence format (0 = percentage, 1= absolute number) and volume difference magnitude (0 = small, 1=large) as the independent variables. Since the interaction between volume difference and valence format was not significant ($\beta = .51$, $p = .46$), I did not include the interaction term in the analysis. The overall model showed a moderately good fit (AIC = 204.53, McFadden's R^2 = .05). As expected, valence format had a significant negative effect on the likelihood of choosing the high-volume option ($\beta = -.86$, $p < .05$). When valence was in absolute numbers (versus percentages), the odds of choosing the high-volume option over the low-volume option was 42% the odds ratio under the percentage format condition. Looking more specifically into participants' actual choices, the high-volume option was selected by 67.86% of participants in the percentage conditions but only 45.71% of participants in the absolute number conditions (χ^2 (1) = 6.79, $p < .001$). The results thus lend support to my first hypothesis and suggest that valence framed in percentages (vs. absolute numbers) increases consumers' likelihood of choosing the higher-volume option.

The logistic regression also revealed a marginally positive effect of volume difference magnitude on the likelihood of choosing the high-volume product (β =.39, $p < .06$). Specifically, the high-volume option was more likely to be chosen under the large volume difference conditions than under the small volume difference conditions (66.67% vs. 49.37%, χ^2 (1) = 10.38, $p < .05$). This is not surprising due to the significantly larger advantage of the high-volume product under the large-difference conditions.

DISCUSSION

The first study provides initial support for my claim that the way review valence is framed can affect consumers' choices. Specifically, consumers are less likely to choose the higher-volume option when the valences of the two options are in absolute numbers than when they are in percentages. More importantly, the findings were robust to small or large volume differences between the two options, such that the impact of volume is significantly attenuated by the absolute number format even when the high-volume product has a very large volume advantage.

STUDY 1B: NUMERICAL FRAMING EFFECT AND MINIMAL VALENCE GAP

Study 1b extends the first experiment in three important ways. First, study 1a involved a tradeoff task where the valence of the low volume product was significantly more positive than that of the high volume product, which could have worked in favor of the low volume option especially under the absolute number format. Study 1b provides a more stringent test of the hypothesis by having the same recommendation rate (80%) for both products. Second, in this study, participants were exposed to the two product options sequentially rather than simultaneously. This sequential exposure better reflects real life situations where consumers usually engage in sequential search and viewing of products. More importantly, it may make it more difficult for consumers to adopt a piecemeal approach since that would require consumers to compare information between two different pages for each review characteristic. Therefore, if the percentage approach still shows a stronger focus on review volume than the absolute number format, it would suggest the piecemeal approach to be quite persistent. Third, study 1b measured participants' purchase intention along a scale instead of forcing them to choose a single product since some participants may decide to be neutral.

METHODOLOGY

Participants and Procedures

Since volume difference magnitude did not interfere with the effects of valence format as shown in the last study, I developed the current study with only two conditions – percentage and absolute number formats. Same as study 1a, participants were asked to examine the online reviews of two pairs of shoes with the exception that the information for each product was displayed on its own page instead of together. The order in which the two products were displayed was randomized. One product had 54 total reviews, and the other had 214 reviews, while the valence was both 80% (or 43 and 174 consumers respectively recommend in the absolute number format). After examining the information of the two products, participants indicated their purchase intention on a 100-point semantic scale with the two product options as anchors, where 0 means they would definitely choose the low volume option and 100 means that they would definitely choose the high volume option. I also recorded participants' demographic information. The study was conducted with 130 MTurk respondents (mean age = 38.27, 60.00% female,).

RESULTS

I ran an ANOVA with purchase intention as the dependent variable and numerical framing, product exposure order, and their interaction as the independent variables. Neither the interaction ($p = .22$) nor the main effect of exposure order ($p = .33$) was significant, indicating that exposure order did not influence consumers' decisions. Therefore, I left exposure order out of the subsequent analysis. In order to test the main effect of numerical framing, I ran a paired-comparison t-test of purchase intention between the two numerical framing conditions. Consistent with study 1a, the effect was significant ($t = 2.71$, $p <. 01$). The results showed that consumers' likelihood of choosing the high volume product was smaller when valence was in absolute numbers (M = 53.11 out of 100) than when it was in percentages (M = 65.29 out of 100). Thus, hypothesis 1 was supported.

DISCUSSION

The results of this study confirm that the absolute number format indeed helps boost consumer purchases of quality products with low number of online reviews. The results also lend support for the existence of the piecemeal approach when valence is presented as percentages, even when products are not displayed side by side. In the next study, I extend the results to other product categories and to situations where the low volume product has a really low number of reviews.

STUDY 2: NUMERICAL FRAMING EFFECT AND LARGE VOLUME GAP

The purpose of this study is to test the robustness of the results found in the first set of studies in two ways. First of all, studies 1a and 1b only investigated a single product category (shoes). This study extends to two other categories - blankets and microwave ovens. These two new categories are more utilitarian than shoes. It is possible that consumers spend more effort processing the numbers and thus are less likely to use heuristics to make decision. The effect of the absolute number framing thus can potentially be smaller for those product categories. More importantly, in studies 1a and 1b, the low volume option still had a relatively high number of reviews (50 total reviews). This could have helped consumers to choose the low volume product as a review volume of 50 may already be considered sufficient for at least some consumers. Addressing this issue, the low volume product in this study will have just a handful of product reviews.

METHODOLOGY

Participants and Procedures

Study 2 featured a 2 (valence format: absolute number versus percentage) x 3 (products: shoes, blankets, and microwave ovens) between-subjects design. Two hundred and forty nine MTurk respondents participated in this study (46.59% female, average age = 40.91). They were randomly assigned to one of the six conditions. The procedure was the same as studies 1a, except that the low-volume product had 11 reviews with an 82% recommendation rate (9 customers recommend) and the high volume option had 450 reviews with an 80% recommendation rate (360 customers recommend). Respondents then made a binary selection indicating which product they would choose. To check the manipulation, I again asked respondents to indicate whether they agreed that the recommending reviews were presented in absolute numbers rather than in percentages.

RESULTS

To test the manipulation of numerical format, an ANOVA was conducted with recalled framing as the dependent variable and numerical framing, product categories, and their interaction as the independent variables. Numerical framing had significant effect on recalled framing ($F_{1,244} = 97.94$ and $p < .001$). A further contrast analysis showed that the respondents who were in the absolute number conditions were more likely to agree that valence was presented as absolute numbers ($M_{number} = 5.89$) compared with those who were in the percentage format conditions ($M_{percentage} = 3.39$), suggesting successful manipulation. For product categories, its main effect ($F_{1,244} = .66$ and $p = .42$) as well as its interaction with numerical format ($F_{1,244} = .11$ and $p = .73$) were not significant.

Regarding the main hypothesis, I first ran a logistic regression with respondents' choice as the dependent variable and review framing, product categories, and their interaction as the independent variables. Neither the main effect of product categories ($p = .82$) nor the interaction was significant ($p = .63$). Therefore, the two product categories were collapsed in the subsequent analysis. Consistent with the earlier studies, review framing weakened the effect of review volume on consumers' choices (χ^2 (1) = 11.09, $p < .001$). The percentage of participants who chose the high volume option was reduced from 88.49% under the percentage framing to 70.9% under the absolute number framing. Not surprisingly, with a very low number of reviews for the low-volume product, most consumers favored the high-volume option. However, even with a strong disadvantage, the bias against the low volume product was still reduced by the absolute number format.

DISCUSSION

Under more rigorous conditions, study 2 showed that the ability of absolute number (vs. percentage) framing to reduce the dominance of the high volume product still holds. Even when the low volume option has a very limited number of reviews, its minimally more positive valence becomes more salient under the absolute number framing. As a result, consumers' likelihood of choosing that option versus the high volume option increases. Furthermore, study 2 suggests that the effect of framing works for different product categories. One limitation of the studies reported so far is that I did not examine explicitly the underlying processes through which consumers evaluate the review information. This is addressed in the next study using the eye-tracking technique.

STUDY 3: CONSUMERS' PROCESSING OF VALENCE AND VOLUME

This study extends the previous studies by showing the process through which consumers examine valence and volume. Specifically, I tracked the movements of consumers' eyes while they examined the product review information. For both numerical framings, consumers are expected to look at the valence and volume of the two products. But hypothesis 1 and its rationale dictate different ways in which one's eyes attend to such information. Specifically, when valence is represented in a percentage format, the piecemeal processing strategy implies that consumers will move their eyes' focus between the valences of the two products and between the volumes of the two products. Meanwhile, the holistic approach deployed under an absolute number format should make consumers' visual focus more likely to move between the valence and volume information within each product.

METHODOLOGY

Participants and Procedures

Forty business students participated in the study for extra course credits. The study sessions took place individually in a lab setting. Participants first filled out information about their eyes' conditions and gave their consent regarding the anonymous use of their data. Next, they were asked to sit in front of a 14-inch laptop with a Tobii 4C eye tracker installed on the lower border of the computer screen. This eye tracker is able to capture eye movements without requiring users to wear any apparatus, thus enhancing the naturalism of the study. The eye tracker was calibrated for each participant before he or she moved on to the main task. If initial calibration failed, participants adjusted their chair so that the eye tracker could accurately recognize their eye movements. Once calibration was successful, participants were informed that they would be shown the information for two pairs of shoes on the screen and that they would be asked about their decision later.

Similar to study 1a, participants were exposed to the review information of the two products simultaneously. One product had 50 reviews with an 88% recommendation rate (44 customers recommend) and the other had 200 reviews with an 80% recommendation rate (160 customers recommend). Participants were randomly assigned to see the valence information either in percentages or absolute numbers. After respondents examined the two products' review information, they provided their demographic information. Each session took approximately 20 minutes.

RESULTS

I captured participants' gaze data while they examined the product information. Specifically, the x and y pixel coordinates of each gaze point was recorded, along with its corresponding time. Figure 2 shows a visual representation of the gaze data. Using the gaze data, I first calculated each participant's total number of fixations, which are defined as any gaze longer than 60 milliseconds (Cian et al. 2014). I then computed each participant's saccades, which refer to eye movements between fixations. In order to test my hypothesis, I created four areas of interest each representing either the valence or the volume of one of the two product options. From this, I created two variables based on the areas of interests – the number of *inter-product saccades* and the number of *intra-product saccades*. Saccades whose fixations jumped from the valence or volume of one product to the valence or volume of the other product were classified as *inter-*

product saccades. Meanwhile, those saccades where fixations jumped between the valence and volume of the same product were considered *intra-product saccades*. Other saccades that did not fall within the four areas of interest were skipped. This approach was adopted from Pieters and Warlop (1999), who captured inter-brand versus intra-brand saccades to study consumers' visual attention during brand choice.

Insert Figure 2 about here

To test the different underlying processes, I created a new variable as the difference between the number of intra-product saccades and that of inter-product saccades. The higher this saccade difference was, the more a participant would be paying attention to within-product information relative to inter-product comparisons. I ran a t-test of this saccade difference between the two framing conditions. The analysis revealed a significantly effect of numerical framing on the saccade difference variable ($t = -2.11$, $p < .05$). Specifically, participants in the absolute number format condition had on average 7.45 more intra-product saccades than inter-product saccades, whereas participants in the percentage framing condition had a significantly smaller difference of 4.72 between intra-product and inter-product saccade counts. The results thus support a more holistic approach utilized under the absolute number framing than under the percentage framing.

It should be noted that even under the percentage framing condition, participants still had more intra-product saccades than inter-product saccades. This could imply that before consumers took the piecemeal approach, they indeed looked at both the valence and volume of the same product. Since transformation of the recommendation percentage into an absolute number is complicated and deemed unnecessary (Slovic 1972, Weathers et al. 2012), participants then proceeded to use the percentage information as is and applied the piecemeal approach.

DISCUSSION

Whereas the earlier experiments revealed the effect of absolute number (vs. percentage) framing on consumers' choices, study 3 showed the mechanism that leads to such an effect. Specifically, when visually examining the product review information, consumers in the absolute number condition were much more likely to shift their eye attention between information for the same product than between the equivalent information (e.g., valence) for the two products, supporting a holistic approach. In contrast, consumers in the percentage condition showed a much smaller gap between their number of intra-product saccades versus inter-product saccades, suggesting a more piecemeal approach. In the fourth and final study, I explore the underlying process in a different way by introducing another representational factor that may disrupt holistic processing.

STUDY 4: MODERATING EFFECT OF COLOR FORMAT

HYPOTHESIS DEVELOPMENT

Study 3 shows that individuals indeed pay attention to review valence and volume differently depending on the numerical representation. This last study presents another test of the underlying process. If consumers are indeed less likely to choose the high-volume product because the absolute number format triggers holistic processing, interrupting that holistic evaluation process and nudging it toward piecemeal processing instead should reduce the difference between the percentage and absolute number formats. Specifically, study 4 introduces different color representations of valence and volume information to disrupt holistic processing. In practice, firms such as Groupon.com apply color distinction between valence and volume numbers (see Figure 3 for an example). Prior research suggests that color is a strong visual factor that primates use to categorize or sort objects (Olson and Poom 2005, Santos et al. 2001). For example, Santos et al. (2001) show that monkeys distinguish edible objects from others on the basis of color. Similarly, Wilcox (1999) report that infants use colors to classify objects. Following this logic, consumers are likely to expect items in the same color to belong to one category and those in different colors to represent different categories. Therefore, even when both valence and volume are expressed as absolute numbers, the different colors used for the two pieces of information are likely to signal them as two distinct attributes. As a result, consumers will be more inclined to use the piecemeal approach, which is adopted by consumers under the percentage format condition. Hence, when valence and volume are presented in different colors, we should see consumers to behave more similarly to each other between the two numerical formats. This leads to the following hypothesis:

Hypothesis 2: The ability of an absolute number framing to weaken the impact of volume on purchase likelihood will be lower when valence and volume are shown in two different colors than when they are in the same color.

Insert Figure 3 about here

METHODOLOGY

Participants and Procedures

The study featured a 2 (valence format: percentage versus absolute number) x 2 (color representation: same versus different colors for volume and valence) between-subjects design. A hundred and eighty five individuals from Amazon Mechanical Turk participated in the study (50.27% male, average age = 34.3). Participants were randomly assigned to one of the four conditions. The study procedure was the same as study 1a, where participants were exposed to the review information for two pairs of shoes and were asked to choose which one of the two they would be more likely to purchase. In each condition, they were exposed to two products simultaneously, with one product having 50 reviews with an 88% recommendation rate (44 customers recommend) and the other having 200 reviews with an 80% recommendation rate (160 customers recommend). For the same color condition, both valence and volume were in black. Meanwhile, for the different color condition, valence was in green and volume was in orange (Figure 4). These are colors commonly used by businesses to display numeric review information (e.g., eaglecreek.com, Amazon.com, HomeDepot.com). As the valence for both products was in the same green color and both review volumes in the same orange color, this design should encourage consumers to treat the two valence numbers as comparable and similarly the two volume numbers as comparable, leading to a piecemeal instead of holistic approach.

Insert Figure 4 about here

To check the manipulations, I asked participants to indicate the extent to which they agree that the recommendation numbers were presented in raw numbers rather than in percentages. Participants also indicated the extent to which they agree that the review numbers they saw were displayed in colors. I further included one item measuring perceived realism of the choice situation by asking participants whether they could imagine the scenario happening to them. All of these items were on a seven-point scale anchored at strongly disagree and strongly agree.

RESULTS

Overall, participants believed that the scenario was realistic ($M = 5.78$). To check the numerical format manipulation, I ran an ANOVA with recalled numerical format as the dependent variable, and numerical format condition, color, and their interaction as the independent variables. The only significant effect from the analysis was numerical format condition ($F_{1,181} = 49.03$ and $p < .001$). Those who were exposed to the absolute number format were more likely to agree that the recommendation information was presented in absolute number ($M_{number} = 5.52$ vs. $M_{percentage} = 3.56$). The main effect ($F_{1,181} = .84$ and $p = .36$) and interaction of color format ($F_{1,181} = .11$ and $p = .74$) were insignificant. A similar ANOVA was conducted with recalled color as the dependent variable. Only color condition had a significant effect such that respondents who were in the different colors condition were more likely to agree that the review information was displayed in colors, compared with those exposed to the same color scenario ($M_{same\ color} = 4.49$ vs. $M_{different\_colors} = 5.71$; $F_{1,180} = 40.41$ and $p < .001$). Numerical framing did not have any main effect ($F_{1,180} = 1.43$ and $p = .23$) or interaction effect with color format ($F_{1,180} = .25$ and $p = .62$) in recalled color.

Hypothesis 2 states that the effect of numerical framing should weaken when valence and volume are presented in different colors. To test this hypothesis, I ran a logistic regression with purchase choice as the dependent variable (1=choosing the high-volume item and 0 otherwise), and numerical format, color, and their interaction as the independent variables. The model had a moderate fit (AIC = 248.18, McFadden's $R^2 = .04$) and showed a significant interaction between numeric format and color ($\beta = 1.34$, $p < .05$). Two chi-squared tests were conducted to examine the effect of numeric format on consumers' decision under different color condition. Specifically, the results showed that if valence and volume were in the same color, consumers' choice of the high volume product was significantly lower when valence was in the absolute number format (26.53%) than when it was in the percentage format (58.33%) (χ^2 (1) = 8.79, $p <.01$) (Figure 5). This replicates the earlier studies' findings. In contrast, when valence and volume were in different colors, the choice share of the high-volume product did not differ between the two numeric format conditions (39.6% for the absolute number condition versus 40% for the percentage condition, χ^2 (1) = .00, $p = 1$). Hypothesis 2 was supported.

Insert Figure 5 about here

DISCUSSION

The purpose of study 4 was to further verify the piecemeal versus holistic approach consumers are likely to adopt when processing absolute number versus percentage valence information. Specifically, the second hypothesis presents color as a boundary condition to the effect of numerical framing. As individuals are more likely to process things of the same colors together and things of different colors apart, when the two valence/volume numbers are in the same color but the valence color and volume color differ, consumers no longer engage in holistic processing even in the absolute number format but instead switch to piecemeal processing. Consequently, the choice outcome becomes similar between the two numerical formats.

GENERAL DISCUSSION

Through five lab experiments, the current research examines how the numerical framing of review valence and volume information in a binary rating system (e.g., thumbs up/down) can affect consumers' processing of such information and their subsequent choice between products. Studies 1 and 2 show that expressing valence as an absolute number leads to a higher choice share for the low-volume product than when valence is expressed as a percentage of the total review volume (percentage of individuals recommend). Tracking of consumers' eye movements in Study 3 shows that the effect of numerical framing observed in earlier studies is due to differences in processing strategy. The absolute number representation triggers a holistic approach, where the same information type is first compared between products (e.g., valence for product A versus valence for product B) and then compared between information types (e.g., valence difference versus volume difference). In contrast, the percentage representation triggers a piecemeal approach that involves comparing the valence and volume information for the same product first to identify the gap between the two, which is then compared between products. Finally in Study 4, I show that color serves as a boundary condition to the effects above such that presenting valence and volume information in different colors disrupts holistic processing in favor of a piecemeal approach and consequently erase the effect of numerical framing.

THEORETICAL IMPLICATIONS

This research provides three important contributions to the marketing literature. First, although the impact of valence and volume on consumers' decisions has been a prominent topic in online word-of-mouth research, *how* consumers process these two pieces of information is still under researched (King et al. 2014). Prior studies have drawn conflicting conclusions especially with regard to the effect of review volume. While some research shows an important impact from online review volume (e.g., Dellarocas, 2003; Liu, 2006), others find volume as not that impactful (Chintaguna et al., 2010). Some of these discrepancies can be better understood within the context of the psychological mechanisms through which consumers pay attention to and interpret review volume and valence information. Specifically, the current research suggests that the effect of volume and valence information is not static. Rather, it is dependent on how such information is presented to consumers numerically. The framing of review valence and volume numbers can alter consumers' approach to interpreting such information and subsequently change their purchase decisions.

Second, previous studies on numerosity tend to assume that consumers utilize a holistic approach and consider all numeric information of one product before comparing between products. Thus, research participants in these studies were typically exposed to only one product before making their decisions (e.g., Weathers et al. 2012). This prevents researchers from exploring the possibility that the presence of more than one option may alter how consumers respond to different numeric information. Using paired comparison choice tasks, the current research shows that such multi-option decision tasks give consumers the opportunity to adopt completely different processing strategies depending on the numeric framing of within-product information. These findings reveal the malleable nature of numerosity effects contingent on the decision context. In future numerosity studies, researchers may want to consider incorporating different types of decision tasks to better identify consumers' thought processes.

Finally, color is usually considered an aesthetic factor (Labrecque and Milne 2012). It has received limited attention from both the numerosity literature and the online review literature. In the numerosity literature, prior research has typically manipulated only numeric information such as how much discount a product offers or how such discount information is presented numerically (e.g., DelVecchio et al. 2007). Similarly, online review studies have typically focused on how the numeric changes in valence or volume influence consumer decisions (e.g.., Liu 2006). My research shows that color can play an important role in consumers' processing of numbers. Specifically, consumers are likely to group numeric information based on the colors attached to that information. If review valence and volume are in different colors, consumers will consider them as two distinct attributes even when both valence and volume are expressed as absolute numbers. As a result, they will adopt a cross-product piecemeal approach instead of a within-product first holistic processing approach.

MANAGERIAL IMPLICATIONS

In an era where online reviews are considered heavily in consumers' purchase decisions, newly introduced products with a low number of reviews are at a significant disadvantage compared with incumbent products that may have accumulated many reviews, even if the new product is superior in quality. As a result, marketers of new products often solicit online reviews from consumers by all means necessary. Specifically, retailers are frequently willing to offer free products in exchange for positive consumer reviews, especially five-star ones (Conger 2016). This practice undoubtedly creates biased reviews that can hurt the general helpfulness of online reviews and consumer welfare (Chevalier and Mayzlin 2006). Therefore, major online retailers such as Amazon.com have banned their vendors from soliciting online reviews through free product offerings (Perez 2016). Although well intentioned, such a policy makes it even harder for small vendors and new products to compete. This paper suggests another viable approach to solving the problem. By adjusting the representational format of review valence and volume numbers, retailers can help remove the low review volume disadvantage of new products. Accordingly, new products will have a better chance of being considered by consumers.

LIMITATIONS AND FUTURE RESEARCH

This paper did not consider individual differences in consumers' selection of numeric processing approaches. For example, it is assumed that in the absolute number format, consumers will engage in a subtraction exercise instead of converting the absolute number of recommending reviews into a percentage. This may not be true for all consumers, especially those who are highly proficient in arithmetic. Future research should consider possible individual heterogeneity in processing strategies as a result of factors such as numeral literacy. In the meantime, given the fact that real-world valence and volume numbers are usually odd rather than even and consumers often deal with many more than just two products, the likelihood of using the heuristics suggested in this paper could be even higher in reality.

In addition, the current research examined only a binary rating system (i.e., whether consumers recommend or not). It did not consider the more granular multi-score rating systems, as rating dispersion within the same product's reviews can have an effect on consumers' decisions (He and Bond 2015) but is out of the scope of this paper. For such a multi-score rating system (e.g., a five-star rating system), some businesses such as Amazon.com use a percentage format to represent rating dispersion (i.e., what percentage of total reviews is one-star, two-star, etc.), whereas others such as HomeDepot.com show rating dispersion in an absolute number format (i.e., how many reviews are one-star, two-star, etc.). When percentage framing is used, the large number of one-star reviews of a high-volume product can become relatively small, compared with low-volume products. Future research should examine how numerical framing of rating dispersion in the context of a graduated rating system as an extension to the current findings.

Finally, this paper did not consider retailer and product factors that may interfere with the effect of numerical framing. For example, prior research suggests that luxury hotels can overcome the problem of low review volume (Blal and Sturman 2014). Particularly, for those luxury hotels, consumers rely more on valence than on volume. Therefore, it is possible that the use of absolute number framing is even stronger for those luxury brands. Hence, it is likely that brands play an important role in the process. The effects of such retailer and product factors can be another direction that warrants future research attention.

REFERENCES

Anderson, N. H. (1972). Information Integration Theory: A Brief Survey. *Contemporary Developments in Mathematical Psychology, 2*.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323.

Blal, I., & Sturman, M. C. (2014). The differential effects of the quality and quantity of online reviews on hotel room sales. *Cornell Hospitality Quarterly*, *55*(4), 365-375.

Cadario, R. (2015). The impact of online word-of-mouth on television show viewership: An inverted U-shaped temporal dynamic. *Marketing Letters*, *26*(4), 411-422.

Chen, S. F. S., Monroe, K. B., & Lou, Y. C. (1998). The effects of framing price promotion messages on consumers' perceptions and purchase intentions. *Journal of Retailing*, *74*(3), 353-372.

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*(5), 944-957.

Cian, L., Krishna, A., & Elder, R. S. (2014). This logo moves me: Dynamic imagery from static images. *Journal of Marketing Research*, *51*(2), 184-197.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407-1424.

DelVecchio, D., Krishnan, H. S., & Smith, D. C. (2007). Cents or percent? The effects of promotion framing on price expectations and choice. *Journal of Marketing*, *71*(3), 158-170.

DeMers, J. (2015, December 28). How Important Are Customer Reviews For Online Marketing? Retrieved on June 09, 2017 from: https://www.forbes.com/sites/jaysondemers/2015/12/28/how-important-are-customer-reviews-for-online-marketing/#5f1aff101928

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

Fiske, S. T. (1982). Schema-triggered affect: Applications to social perception. In *Affect and cognition: 17th Annual Carnegie Mellon symposium on cognition* (pp. 55-78). Hillsdale: Lawrence Erlbaum.

Gupta, P., & Harris, J. (2010). How e-WOM recommendations influence product consideration and quality of choice: A motivation to process information perspective. *Journal of Business Research*, *63*(9), 1041-1049.

He, S. X., & Bond, S. D. (2015). Why is the crowd divided? Attribution for dispersion in online word of mouth. *Journal of Consumer Research*, *41*(6), 1509-1527.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing*, *28*(3), 167-183.

Kleber, J., Kleber, J., Florack, A., Florack, A., Chladek, A., & Chladek, A. (2016). How to present donations: the moderating role of numeracy in cause-related marketing. *Journal of Consumer Marketing*, *33*(3), 153-161.

Labrecque, L. I., & Milne, G. R. (2012). Exciting red and competent blue: the importance of color in marketing. *Journal of the Academy of Marketing Science*, *40*(5), 711-727.

Lee, Y. J., Hosanagar, K., & Tan, Y. (2015). Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Science*, *61*(9), 2241-2258.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74-89.

Morwitz, V. G., Greenleaf, E. A., & Johnson, E. J. (1998). Divide and prosper: consumers' reactions to partitioned prices. *Journal of Marketing Research*, 453-463.

Muthukrishnan, A. V., Warlop, L., & Alba, J. W. (2001). The piecemeal approach to comparative advertising. *Marketing Letters*, *12*(1), 63-73.

O'Donovan (2017, April 11). That Four-Star Rating You Left Could Cost Your Uber Driver Her Job. Retrieved June 09, 2017, from https://www.buzzfeed.com/carolineodonovan/the-fault-in-five-stars?utm_term=.oy7KDD5v1#.yn1xKKb7z

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(24), 8776-8780.

Pelham, B. W., Sumarta, T. T., & Myaskovsky, L. (1994). The easy path from many to much: The numerosity heuristic. *Cognitive Psychology*, *26*(2), 103-133.

Perez, S. (2016, October 3rd). Amazon bans incentivized reviews tied to free or discounted products. Retrieved April 21, 2018, from: https://techcrunch.com/2016/10/03/amazon-bans-incentivized-reviews-tied-to-free-or-discounted-products/

Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing*, *16*(1), 1-16.

Punj, G. (2012). Income effects on relative importance of two online purchase goals: Saving time versus saving money? *Journal of Business Research*, *65*(5), 634-640.

Santos, L. R., Hauser, M. D., & Spelke, E. S. (2001). Recognition and categorization of biologically significant objects by rhesus monkeys (Macaca mulatta): the domain of food. *Cognition*, *82*(2), 127-155.

Seiler, M., Madhavan, P., & Liechty, M. (2012). Toward an understanding of real estate homebuyer internet search behavior: an application of ocular tracking technology. *Journal of Real Estate Research*.

Slovic, P. (1972). Psychological study of human judgment: Implications for investment decision making. *The Journal of Finance*, *27*(4), 779-799.

Sujan, M. (1985). Consumer knowledge: Effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research*, 31-46.

Tangari, A. H., Burton, S., & Davis, C. (2014). Do They Have Your Number? Understanding the Moderating Role of Format Effects and Consumer Numeracy for Quantitative Front‐ of‐ Package Nutrition Claims. *Journal of Consumer Affairs*, *48*(3), 620-633.

Tao, T., Wyer Jr, R. S., & Zheng, Y. (2017). The role of categorization and scale endpoint comparisons in numerical information processing: A two-process model. *Journal of Experimental Psychology: General*, *146*(3), 409.

Weathers, D., Swain, S. D., & Carlson, J. P. (2012). Why consumers respond differently to absolute versus percentage descriptions of quantities. *Marketing Letters*, *23*(4), 943-957.

Wilcox, T. (1999). Object individuation: Infants’ use of shape, size, pattern, and color. *Cognition*, *72*(2), 125-166.

Yang, J., Kim, W., Amblee, N., & Jeong, J. (2012). The heterogeneous effect of WOM on product sales: why the effect of WOM valence is mixed? *European Journal of Marketing*, *46*(11/12), 1523-1538.

You, Y., Vadakkepatt, G. G., & Joshi, A. M. (2015). A meta-analysis of electronic word-of-mouth elasticity. *Journal of Marketing*, *79*(2), 19-39.

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

Santos, L. R., Hauser, M. D., & Spelke, E. S. (2001). Recognition and categorization of biologically significant objects by rhesus monkeys (Macaca mulatta): the domain of food. *Cognition*, *82*(2), 127-155.

FIGURES

FIGURE 1: STUDY 1 – AN EXAMPLE OF STIMULI

FIGURE 2: STUDY 3 – EYE-TRACKING DATA VISUALIZATION

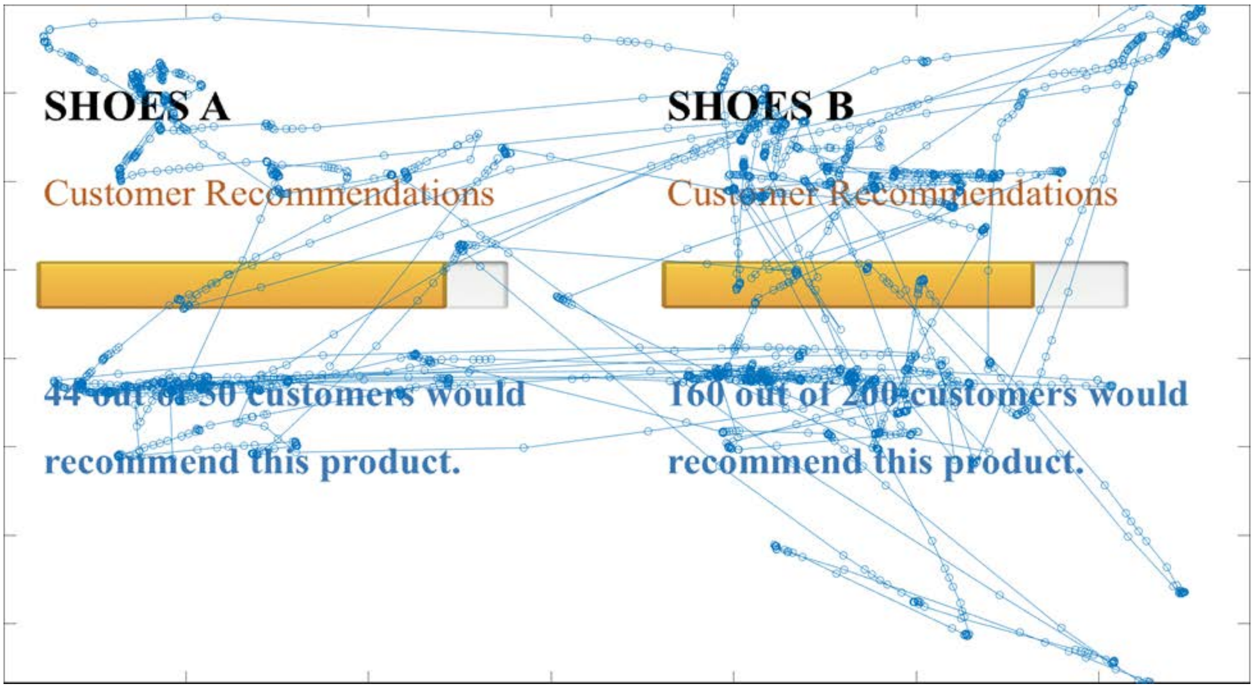


FIGURE 3: BINARY REVIEW SYSTEM FROM GROUPON.COM SITE

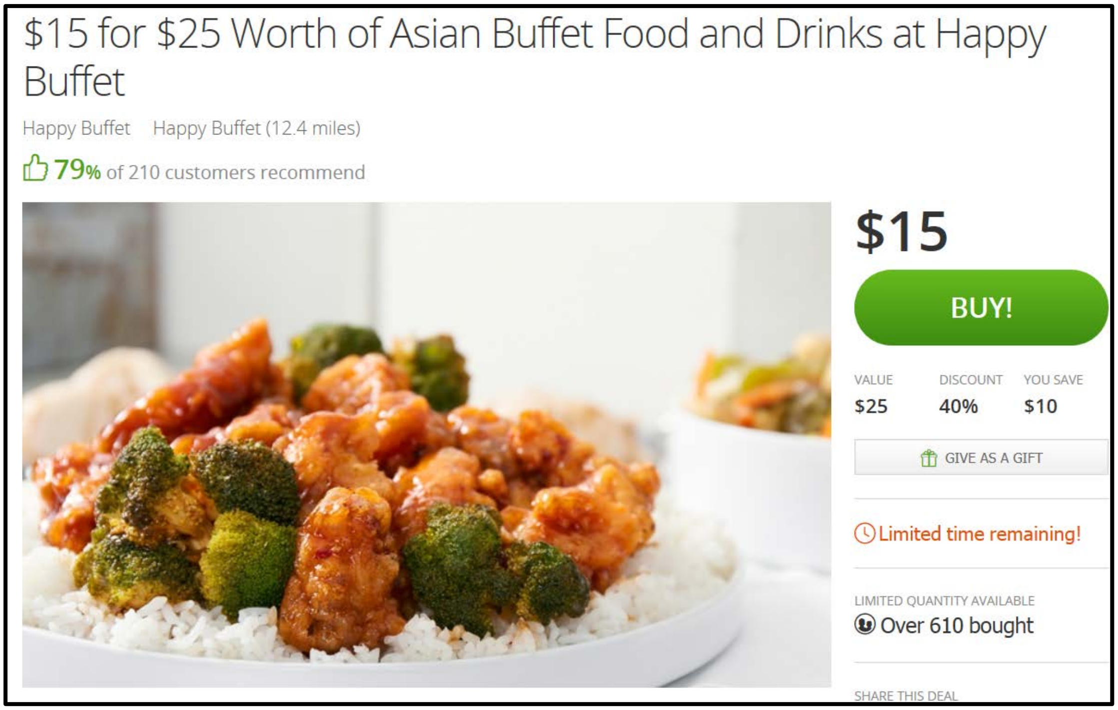


FIGURE 4: STUDY 4 - STIMULUS

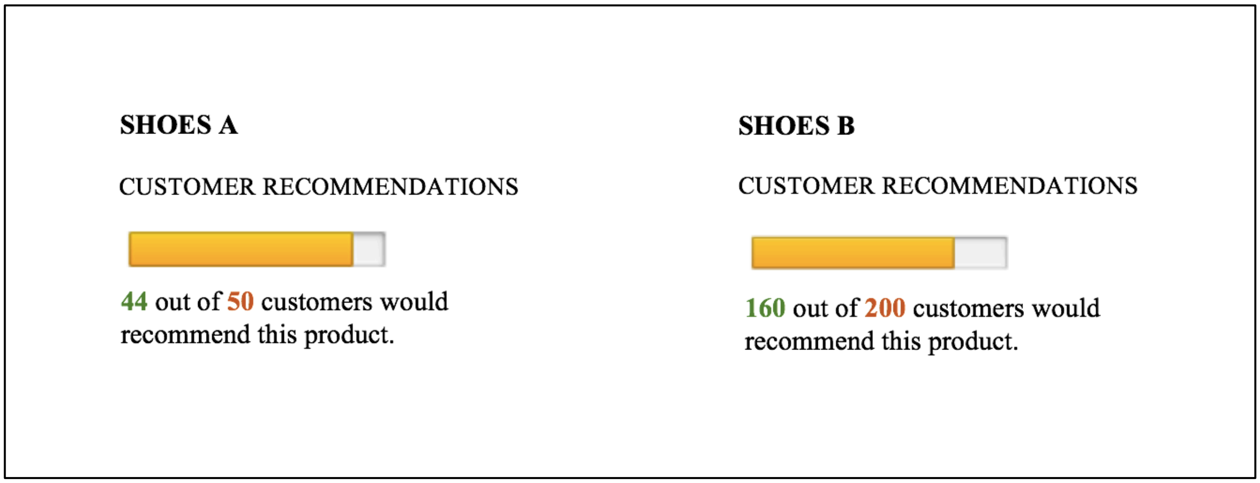
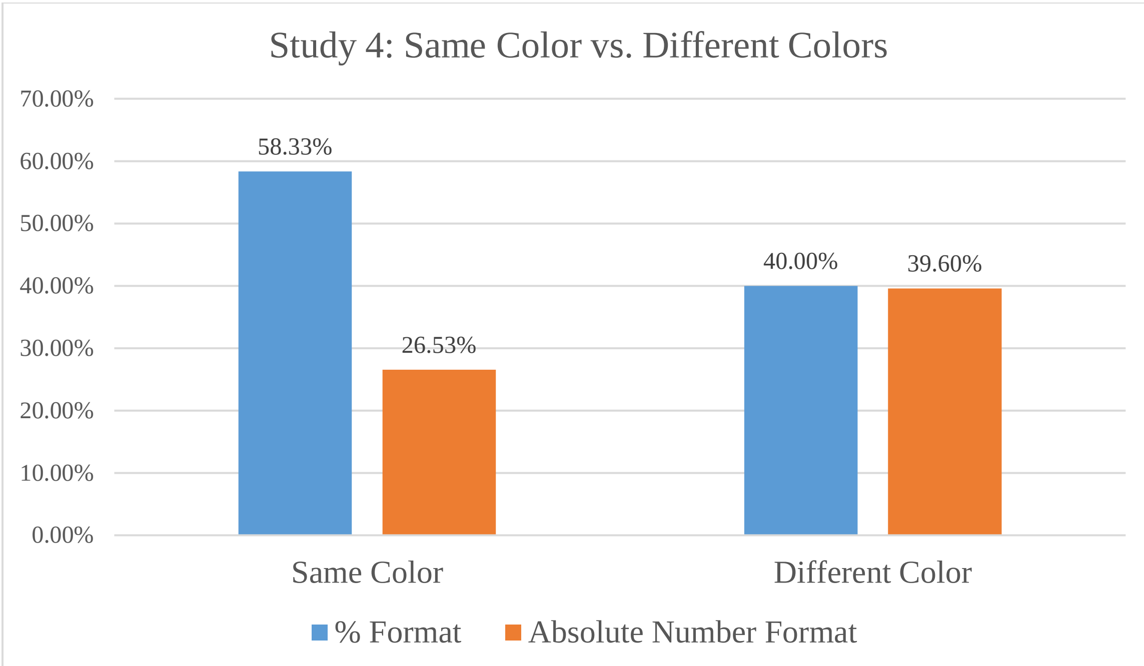


FIGURE 5: STUDY 4 - RESULTS



VITA

ANH DANG
Strome College of Business, Old Dominion University
2160 Constant Hall, Norfolk, VA 23529

EDUCATION

2014-2018 PhD in Business Administration – Marketing, Old Dominion University, Norfolk, VA

2011-2013 MBA, University of Houston-Clear Lake, Houston, Texas

2006-2010 BS in Business Administration – Marketing, International University, Ho Chi Minh National University of Vietnam, Vietnam.

JOURNAL PUBLICATIONS

Xu, Hangjun, Huiling Guo, Jing Zhang, & **Anh Dang** (2018). Facilitating Dynamic Marketing Capabilities Development for Domestic and Foreign Firms in an Emerging Economy. *Journal of Business Research*.

Dang, Anh & Aaron Arndt (2017). How Personal Costs Influence Customer Citizenship Behaviors. *Journal of Retailing and Consumer Services*.

SELECT CONFERENCE PROCEEDINGS

Dang, Anh, & Hangjun Xu (2017). Who Does the Good Deed? The Effect of Message Framing on Green Consumption. *AMA Summer Educators' Conference*, San Francisco, CA.

Nguyen, Trung, & **Anh Dang**. Corporate Social Responsibility and Market Reactions: The Role of CEO Risk Preferences. *Financial Management Association's Conference*, Boston, MA (2017).

Kirande, Kiran, **Anh Dang,** Olga Butenko Moseley, & Hangjun Xu (2017). Let's Pin It Down: The Role of Displayed Donations in Corporate Philanthropy. *AMA Winter Educators' Conference*, Orlando, FL.

Dang, Anh, & Girish Sreevatsan Nandakumar (2016). Cultural Values and Their Impact on Electronic Word of Mouth (eWOM). *AMS World Marketing Congress Conference*, Paris.

Dang, Anh (2015). Hedonic vs. Utilitarian Products: The Dawn of Intra-Product Category Research. *AMS Conference*, Denver, CO.

Dang, Anh, Olga Butenko Moseley, & Hangjun Xu (2015). Conditional Corporate Philanthropy: Target Markets and Marketing Mix. *AMA Summer Educators' Conference*, Chicago, IL.