Exploring the Role of Preference Heterogeneity and Causal Attribution in Online Ratings Dynamics*

Wujin Chu**
Minjung Roh***

This study investigates when and how disagreements in online customer ratings prompt more favorable product evaluations. Among the three metrics of volume, valence, and variance that feature in the research on online customer ratings, volume and valence have exhibited consistently positive patterns in their effects on product sales or evaluations (e.g., Dellarocas, Zhang, and Awad 2007; Liu 2006). Ratings variance, or the degree of disagreement among reviewers, however, has shown rather mixed results, with some studies reporting positive effects on product sales (e.g., Clement, Proppe, and Rott 2007) while others finding negative effects on product evaluations (e.g., Zhu and Zhang 2010). This study aims to resolve these contradictory findings by introducing preference heterogeneity as a possible moderator and causal attribution as a mediator to account for the moderating effect.

The main proposition of this study is that when preference heterogeneity is perceived as high, a disagreement in ratings is attributed more to reviewers' different preferences than to unreliable product quality, which in turn prompts better quality evaluations of a product. Because disagreements mostly result from differences in reviewers' tastes or the low reliability of a product's quality (Mizerski 1982; Sen and Lerman 2007), a greater level of attribution to reviewer tastes can mitigate the negative effect of disagreement on product evaluations. Specifically, if consumers infer that reviewers' heterogeneous preferences result in subjectively different experiences and thereby highly diverse ratings, they would not disregard the overall quality of a product. However, if consumers infer that reviewers' preferences are quite homogeneous and thus the low reliability of the product quality contributes to such disagreements, they would discount the overall product quality. Therefore, consumers would respond more favorably to disagreements in ratings when preference heterogeneity is perceived as high rather than low.

This study furthermore extends this prediction to the various levels of average ratings. The heuristic-systematic processing model so far indicates that the engagement in effortful systematic processing occurs only when sufficient motivation is present (Hann et al. 2007; Maheswaran and Chaiken 1991;

** Professor, Graduate School of Business, Seoul National University(wchu@snu.ac.kr)
*** Ph. D. Candidate, Graduate School of Business, Seoul National University(pagani11@naver.com), Corresponding author

Martin and Davies 1998). One of the key factors affecting this motivation is the aspiration level of the decision maker. Only under conditions that meet or exceed his aspiration level does he tend to engage in systematic processing (Patzelt and Shepherd 2008; Stephanous and Sage 1987). Therefore, systematic causal attribution processing regarding ratings variance is likely more activated when the average rating is high enough to meet the aspiration level than when it is too low to meet it. Considering that the interaction between ratings variance and preference heterogeneity occurs through the mediation of causal attribution, this greater activation of causal attribution in high versus low average ratings would lead to more pronounced interaction between ratings variance and preference heterogeneity in high versus low average ratings. Overall, this study proposes that the interaction between ratings variance and preference heterogeneity is more pronounced when the average rating is high as compared to when it is low.

Two laboratory studies lend support to these predictions. Study 1 reveals that participants exposed to a high-preference heterogeneity book title (i.e., a novel) attributed disagreement in ratings more to reviewers' tastes, and thereby more favorably evaluated books with such ratings, compared to those exposed to a low-preference heterogeneity title (i.e., an English listening practice book). Study 2 then extended these findings to the various levels of average ratings and found that this greater preference for disagreement options under high preference heterogeneity is more pronounced when the average rating is high compared to when it is low.

This study makes an important theoretical contribution to the online customer ratings literature by showing that preference heterogeneity serves as a key moderator of the effect of ratings variance on product evaluations and that causal attribution acts as a mediator of this moderation effect. A more comprehensive picture of the interplay among ratings variance, preference heterogeneity, and average ratings is also provided by revealing that the interaction between ratings variance and preference heterogeneity varies as a function of the average rating. In addition, this work provides some significant managerial implications for marketers in terms of how they manage word of mouth. Because a lack of consensus creates some uncertainty and anxiety over the given information, consumers experience a psychological burden regarding their choice of a product when ratings show disagreement. The results of this study offer a way to address this problem. By explicitly clarifying that there are many more differences in tastes among reviewers than expected, marketers can allow consumers to speculate that differing tastes of reviewers rather than an uncertain or poor product quality contribute to such conflicts in ratings. Thus, when fierce disagreements are observed in the WOM arena, marketers are advised to communicate to consumers that diverse, rather than uniform, tastes govern reviews and evaluations of products.

Key words: electronic word-of-mouth (eWOM), online customer ratings, causal attribution, disagreement, preference heterogeneity, aspiration level

Ⅰ. Introduction

Spending money involves a procedure as cognitively demanding and effortful as earning money. The latest reliable information sources are often required to guarantee a good consumption outcome from this procedure. Consider, for example, an office worker who eagerly looks forward to his annual vacation. He may have dreamed of an extraordinary backpacking trip abroad. In order not to spoil his holiday time, he may have referred to various information sources, such as travel agencies' websites, professional travelers' blogs, or online travel communities that contain the advice of peer customers. Among these sources, customers' online reviews occupy an exceptional place in that they engage consumers in a true two-way and meaningful dialogue with peer consumers rather than pushing messages out to them in a one-way flow (Keller 2007). By chatting with and exchanging messages with peer consumers, the office worker may indeed conjure up more vivid and lifelike images about the vacation place, thereby improving the quality of the outcome of his decisions regarding his vacation.

In fact, this type of consumer-to-consumer online communication, or electronic word of mouth (eWOM), has recently received much academic and practitioner attention (Cheng and Zhou 2010). As advertising and editorial content began to decline in importance, word of mouth

(WOM) assumed ever greater significance as an effective communication channel (Keller 2007). For instance, among U.S. consumers, the average daily participation in WOM conversations amounts to 121 times and the proportion of the online population who read consumer-generated feedback approaches 25% (Keller 2007; Li and Bernoff 2008). Among Korean consumers, the average proportion of the online population who post online reviews is as high as 73%, and the proportion of those who read the online feedback of news articles is close to 80% (ETnews.com 2005; Mediatoday 2006). Moreover, a growing body of research has suggested that this ubiquitous usage of eWOM and online reviews is not a mere passing phenomenon but is rather a permanent feature of electronic commerce as it has an enduring effect on product sales, sales growth, and consumer behavior (e.g., Cui, Lui, and Guo 2012; Lin, Luarn, and Huang 2005; Zhu and Zhang 2010).

To date, research regarding online customer ratings has focused on three metrics: the ratings volume, valence, and variance. The findings have demonstrated that the ratings volume exerts a positive impact on product sales and consumer purchase intentions (Duan, Gu, and Whinston 2008; Liu 2006), and that ratings valence (or average rating) positively affects product sales, sales growth, and consumer product evaluations (Chevalier and Mayzlin 2006; Dellarocas, Zhang, and Awad 2007). However,

regarding the variance of ratings, or the extent of disagreement of opinion, the findings are mixed, showing that greater disagreement leads to higher sales in some studies (Clement, Proppe, and Rott 2007; Karniouchina 2011) but not in others (Godes and Silva 2012; Zhu and Zhang 2010).

Thus, to reconcile these contradictory findings, we introduce preference heterogeneity as a possible moderator (Price, Feick, and Higie 1989) and causal attribution as a mediator that underlies this moderation effect (Mizerski 1982; Ryu, Park, and Feick 2006; Sen and Lerman 2007). We propose that when preference heterogeneity is perceived as high, disagreement in ratings is attributed more to the reviewers' different preferences than to the low reliability of the product quality, which in turn, prompts better quality evaluations of products as compared to when preference heterogeneity is perceived as low. In fact, disagreements arise from either the low reliability of product quality or from the different consumption experiences of each reviewer. Products may yield arbitrarily different outcomes across various reviewers due to their low reliability of quality performance, or elicit subjectively different evaluations from each reviewer due to the diverse preferences of the reviewers. Whether the preferences vary across reviewers may indeed direct consumers' interpretations of disagreements in ratings. Thus, we expect that high perceived preference heterogeneity leads to greater attri-

bution to reviewers' tastes for disagreement and thereby prompts more favorable product evaluations with such ratings, as compared to low perceived preference heterogeneity.

Furthermore, we extend these predictions to various levels of average ratings. The heuristic-systematic processing model suggests that the engagement in effortful systematic processing occurs only when sufficient motivation is present (Hann et al. 2007; Martin and Davies 1998). One possible factor that affects this motivation is the aspiration levels of the decision maker; engagement in systematic processing emerges only under conditions that meet or exceed individuals' aspiration levels (Patzelt and Shepherd 2008; Stephanous and Sage 1987). Thus, based on these studies, we suggest that if the average rating seems high enough to meet the aspiration level, the systematic causal attribution processing regarding ratings variance is more activated, and thereby its interaction with preference heterogeneity is more pronounced. However, if the average rating seems too low to meet such aspiration levels, the causal attribution processing regarding ratings variance is less activated; thus, its interaction with preference heterogeneity is rather attenuated. Overall, we expect that the proposed interaction between ratings variance and preference heterogeneity is more pronounced when the average rating is high as opposed to when it is low.

We organize the rest of this paper as follows. First, we review the literature on online rat-

ings, ratings variance, preference heterogeneity, causal attribution, and aspiration levels. Next, we develop the hypotheses to be tested in two studies. In Study 1, we test our key prediction regarding the interaction between ratings variance and preference heterogeneity, and the mediating effect of causal attribution in such interaction, using a novel and an English listening practice book as the focal products. In Study 2, we extend this investigation to the various levels of average ratings to present a more comprehensive picture of these ratings dynamics, adopting a laptop computer as the focal product. Finally, we conclude with a discussion on the implications and the limitations of our research.

II. Theoretical Background and Hypothesis Development

2.1 Online Customer Ratings

Research regarding online customer ratings has so far substantiated their effects on product sales and consumer decisions using the data from retailer websites or consumer review sites. These studies have demonstrated that an improvement in customer ratings positively affects book sales (Chevalier and Mayzlin 2006) and movie sales (Duan et al. 2008) and that incorporating such reviews helps enhance mod-

els' levels of accuracy in predicting box-office revenues (Liu 2006) and television viewership (Godes and Mayzlin 2004).

Attempts to specify boundary conditions that affect online ratings processing also appear in several works. The literature focusing on *product type* indicated that the impact of online reviews is greater for experience (versus search) goods because potential consumers have the opportunity to interact with those goods before purchasing them (Huang, Lurie, and Mitra 2009; Park and Lee 2009). Studies examining the significance of the *message sender* then revealed that engagement in e-WOM is driven by diverse motivations such as self-enhancement, social benefits, and advice seeking (Hennig-Thurau et al. 2004) and that these motivations positively contribute to message clarity, informativeness, and reliability (Yap, Soetarto, and Sweeney 2013). Moreover, research focusing on the *message receiver's* characteristics showed that consumers' familiarity with a retailer mitigates their sensitivity to negative e-WOM (Chatterjee 2001), whereas greater Internet experience sharpens such sensitivity to negative e-WOM (Zhu and Zhang 2010). Finally, research examining the *message characteristics* revealed that the length of the e-WOM communication exerts a significant effect on purchase intentions (Lin et al. 2005) and that face-to-face WOM communications is more persuasive than that in printed form, especially when a prior impression of a stimuli product is

not available in memory (Herr, Kardes, and Kim 1991).

2.2 Online Ratings Variance

Another research stream that is related to the study of message characteristics focused on ratings metrics, specifically the ratings volume, valence, and variance. It has thus far found that volume positively affects product sales and consumer purchase intention (Duan et al. 2008; Lin et al. 2005; Liu 2006). Valence also showed a positive influence on product sales, sales growth, and consumer product evaluations (Bae, Shim, and Kim 2010; Chevalier and Mayzlin 2006; Dellarocas et al. 2007). However, this line of research has yielded mixed findings on the variance of ratings; some studies reported positive effects of varied ratings (Clement et al. 2007; Karniouchina 2011; Martin 2008), while others found that they had negative effects (Danescu-Niculescu-Mizil et al. 2009; Dellarocas, Awad, and Zhang 2004; Godes and Silva 2012; Zhu and Zhang 2010).

Basically, ratings variance refers to disagreements among reviewers with regard to product evaluations, reflecting the heterogeneity of consumers' opinions (Moe and Trusov 2011; Sun 2012). A greater amount of disagreement in ratings reflects less consistency in opinions across various reviewers and creates greater uncertainty for consumers making purchase decisions (Zhu and Zhang 2010). Some studies accordingly showed that such great uncertainty accompanying fierce disagreements can negatively affect consumers' evaluations of review ratings (Amazonbook.com; Danescu-Niculescu-Mizil et al. 2009), subsequent ratings (Amazonbook.com; Godes and Silva 2012), product quality (Gamespot.com; Zhu and Zhang 2010), and future product revenue (Yahoo! Movies website; Dellarocas et al. 2004). Overall, these results conform to the "consensus implies correctness" heuristics, considering consensus as a virtue against irreconcilable disagreement (Chaiken and Stangor 1987).

However, some studies have shown opposing results. A study using data on fiction books indicated that critics' disagreements about the quality of a book increase the sales of that book (Clement et al. 2007). Another study using data on movies revealed that ratings disagreements positively affect the volume of WOM, which, expectedly, contributes to box- office success (Karniouchina 2011; Martin 2008). These studies explained that disagreements among reviewers trigger debates concerning the target product, thereby increasing public awareness of the product. A greater volume of WOM, then, can ultimately lead to higher sales, as desired by most marketers (Karniouchina 2011).

A few attempts have been made to reconcile these contradictions by exploring possible moderators. One well-recognized moderator is an aspiration level suggested by West and Broniarczyk (1998). They showed that people prefer there to be disagreement rather than

agreement in ratings when average ratings fall short of their aspiration levels, but reverse their preferences when average ratings exceed their aspiration levels (Kahneman and Tversky 1979). Another possible moderator that can be taken into consideration is the degree of posting activity, which Moe and Schweidel (2012) demonstrated using the data collected from a national retailer (bazaarvoice.com). Their main findings indicated that the higher the ratings variance, the more positive the subsequent ratings are for less active posters but the more negative such ratings are for more active posters.

Related to this, the interaction between ratings variance and product type has also been noted in some literature. Martin (2008) revealed that people prefer high variance options over low variance options when they choose between two desserts, but they reverse this preference when they choose between two disgusting foods from the "Fear Factor" TV program. Park and Han (2008) also showed that higher ratings variance prompts less favorable product evaluations of search products but more favorable product evaluations of experience products, particularly when the prior product attitude is positive. Here the noteworthy point is that across the two studies, the product stimuli leading to the more favorable responses (i.e., dessert and experience products) were both more greatly affected by individuals' unique tastes than their counterpart stimuli were. Actually, the decision regarding what dessert to have mainly depends on one's own dietary preferences and eating tastes (Freeland Graves and Nitzke 2002; Thomas 2007); the evaluation of experience products also conforms more to subjective judgment and consumers' heterogeneous tastes (Huang et al. 2009; Wright and Lynch 1995). However, the product stimuli leading to less favorable responses (i.e., search products and "Fear Factor" foods) seem to hinge more on broadly shared preferences, in that an evaluation of a search product relies more on an objective judgment process and concrete cues (Huang et al. 2009; Park and Han 2008), and the basic emotions of fear and disgust are universal, and thus the stimuli arousing such emotions are likely to prompt more general disgust (Deater-Deckard, Dodge, and Sorbring 2005; Wicker et al. 2003).

Accordingly, based on these findings, we propose another moderator that may affect the relationship between ratings variance and product evaluations: preference heterogeneity. We speculate that more favorable responses to ratings disagreement will emerge when preference heterogeneity is perceived as high rather than low. Specifically, we presume that greater causal attribution to reviewers' tastes than to product quality contributes to these more favorable responses to ratings disagreement. The next section reviews the literature on preference heterogeneity and causal attribution to formulate these predictions more formally.

2.3 Preference Heterogeneity and Causal Attribution

What one person considers good may not suit the tastes of others. Benefit segmentation serves to address this preference heterogeneity (Feick and Higie 1992). Defined as the extent to which individual tastes and preferences for a good or service vary across consumers, preference heterogeneity helps assess the degree of disagreement in the preferences of various consumers (Price et al. 1989). While high preference heterogeneity indicates substantial variation in consumer preferences and low consensus among various consumers, low preference heterogeneity denotes little variation in consumer tastes and greater agreement among consumers.

This concept of preference heterogeneity helps capture the underlying dimensions that govern consumers' responses to external information such as advertising and WOM. For instance, the effect of source similarity varies as a function of preference heterogeneity such that for products with high preference heterogeneity, source similarity more strongly affects consumers' product evaluations (Feick and Higie 1992) and more positively correlates with consumers' advertising responses (Ryu et al. 2006). Conversely, for products with low preference heterogeneity, the overall product ratings appeal more as shopping aids than do personal recommendations or detailed product specifications, and rating reviews act as better predictors of individual product evaluations than do commentary reviews (He 2012).

In accordance with this research stream, we incorporate the concept of preference heterogeneity into our analysis of online ratings variance, especially in its relevance to causal attribution. Essentially, online customer ratings are the outcomes of the interplay between a product and its reviewers. Thus, when ratings reflect disagreement, consumers can consider two possible reasons for such conflicts: low reliability of the product (*product* or *stimuli attribution*) or different preferences among reviewers (*reviewers* or *non-stimuli attribution*; Calder and Burnkrant, 1977; Mizerski 1982; Sen and Lerman 2007). They may infer that (1) the inherent unreliability of the product's quality drives individual reviewers to have actually dissimilar quality experiences, or (2) reviewers with varied preferences have quite subjectively different feelings about the same product. Hence, when reviewers' different preferences appear as a primary cause for such divergent opinions, the causal attribution of disagreement is likely to be directed more toward the reviewers, or their different tastes. However, when individual tastes are not a critical concern and thereby preference heterogeneity is perceived as low, causal attribution is likely to be directed more toward the product, or its unreliability. Causal attribution, or inferences about the cause of a communicator's generation of the given information, thus would lead consumers to form

different interpretations of disagreements in ratings.

This differential causal attribution due to preference heterogeneity further enables us to predict the direction of the effect of disagreement on product evaluation. Because disagreement is caused either by the low reliability of product quality or by the differences in reviewers' preferences, greater attribution to reviewers' tastes can relieve the negative effect of disagreements on product evaluations. That is, if consumers infer that reviewers' heterogeneous preferences result in subjectively different experiences and thereby highly diverse ratings, then they would not discount the overall quality of a product. However, if consumers infer that reviewers' preferences are quite homogeneous and thus a low level of the reliability of a product's quality contributes to such disagreements, then they would discount the overall product quality. Therefore, when preference heterogeneity is perceived as high, disagreement is more likely to be attributed to reviewers' different tastes, which in turn likely allows a target product to avoid negative perception of its quality. Conversely, when preference heterogeneity is perceived as low, disagreement is more likely to be attributed to low reliability of the product's quality, which in turn, likely has a detrimental effect on product evaluations. We thus expect that consumers' likelihood to evaluate disagreement over agreement options more positively will be higher when preference heterogeneity is perceived as high rather than low. This is stated formally in H1 below.

> *H1: Consumers will more positively evaluate products characterized by ratings disagreement than those characterized by ratings agreement when preference heterogeneity is perceived as high rather than low.*

In brief, we assume that greater attribution to reviewers' tastes for disagreement occurs when preference heterogeneity is perceived as high rather than low. This differential causal attribution then will lead to a more favorable product evaluation with such ratings when preference heterogeneity is perceived as high rather than low. We summarize these arguments in the following hypotheses.

> *H2: Consumers will attribute disagreement in ratings more to reviewers' tastes (versus product quality) when preference heterogeneity is perceived as high rather than low.*
> *H3: These causal attributions will mediate the moderating role of preference heterogeneity in the relationship between ratings variance and product evaluations.*

2.4 Aspiration Levels and Review Information Processing

According to the heuristic-systematic proc-

essing model, people often behave as cognitive misers (Taylor and Fiske 1978). Because people are generally reluctant to engage in effortful systematic processing, they tend to disengage from the further processing of a given piece of information when they feel their motivation levels are not high enough (Martin and Davies 1998). One key motivation factor that drives this processing is the *expectancy* or *aspiration level* of a decision maker (Hann et al. 2007; Kirchler et al. 2009; Maheswaran and Chaiken 1991). When outcomes are unlikely to reach their aspiration levels, people rather reduce their efforts and ultimately disengage from a fruitless information-processing task (Folkman and Moskowitz 2004; Scheir and Carver 2006). The expectation of an undesirable outcome can indeed facilitate a withdrawal from further processing effort (Senko and Hulleman 2013; Sideridis and Kaplan 2011).

It thus seems plausible that whether the average ratings meet or fall short of aspiration levels has some effect on review information processing (Patzelt and Shepherd 2008). If products feature ratings below the aspiration levels, consumers may hesitate to perform further systematic processing due to the low likelihood of achieving a desired outcome (Stephanou and Sage 1987). However, if products receive higher average ratings than the aspiration levels, they may proceed further to effortful systematic processing due to the products' possibly high likelihood of attaining a desired out-

come and being the final choice (Church, Laroche, and Rosenblatt 1985; Klein and Bither 1987; Schindler, Berbaum, and Weinzimer 1987). Therefore, given that a global appreciation of success or failure of alternatives is related to heuristic processing while a detailed causal attribution regarding an outcome is more related to systematic processing (Iglesia 2009; Meyers-Levy and Sternthal 1991; Oliver 1997), we can expect that consumers exposed to low average ratings barely attempt an overall appreciation of products using only such simple heuristic cues of average ratings, while those exposed to high average ratings seek more detailed causal attributions of given ratings by systematically processing the entire ratings metrics.

Note here though that disengagement from this causal attribution process of given ratings can attenuate the interaction between ratings variance and preference heterogeneity because the interaction between ratings variance and preference heterogeneity occurs through the mediation of causal attribution. Engagement in causal attribution processing regarding ratings variance is indeed a prerequisite for an interaction to exist between ratings variance and preference heterogeneity. Therefore, for low-average-ratings alternatives, causal attribution of ratings variance rarely occurs and thus the interaction between ratings variance and preference heterogeneity is likely to be attenuated. Conversely, for high-average-ratings alternatives, the causal attribution processing of ratings var-

iance occurs more often and thus the inter-action between ratings variance and preference heterogeneity is likely more pronounced. Together, because the causal attribution process regarding ratings variance is less activated when the average ratings is low than when it is high, the proposed interaction between ratings variance and heterogeneity may not be so prominent for low-average-ratings alternatives as it may be for high-average-ratings alternatives. We state these predictions more formally in the following hypotheses.

> *H4: The proposed interaction between ratings variance and preference heterogeneity will be more pronounced when the average ratings are high rather than low in terms of (H4a) product evaluations and (H4b) causal attribution.*
>
> *H5: The mediating role of causal attribution in the relationship between ratings variance and product evaluations is more likely to emerge when the average ratings are high rather than low.*

The next two studies were designed to test these predictions. In study 1, we assessed the predictions made in hypotheses 1-3, focusing on the interaction between ratings variance and preference heterogeneity and the mediating role of causal attribution. In Study 2, we replicated the findings of Study 1 and tested the predictions of hypotheses 4-5, concentrat-ing on the interplay among ratings variance, preference heterogeneity, and average ratings.

Ⅲ. Study 1

In Study 1, we examine whether disagree-ment in ratings prompts more favorable prod-uct evaluations under high or low perceived preference heterogeneity (H1). We predict that greater attribution of disagreement to reviewers' tastes leads to more favorable product evalua-tions with such ratings when preference heter-ogeneity is perceived as high rather than low (H2 and H3).

To test this prediction, we adopt a novel and an English listening practice book as the focal product. While the evaluation of a novel differs depending on subjective preferences, the evalu-ation of an English listening practice book hinges more on a shared goal among readers, in this case to improve their English listening skills (Arsego 2009; Erickson 1996). Because novel genres are affected much more by per-sonal tastes and preferences than are foreign language learning books, different readers of the same novel are more likely to voice ex-tremely diverse opinions that range from "very impressive" to "quite terrible," whereas those of a language listening practice book would yield rather similar opinions (Feick and Higie 1992; Holbrook 1999). Thus, given that per-

ceived preference heterogeneity tends to be higher for a novel than for a listening practice book, we can expect that the attribution of disagreement to reviewers' tastes is greater for a novel than it is for a listening book, which in turn leads to a more favorable product evaluation for the novel than for the listening book.

Also, to control for the potential confounding effects of individual differences, we additionally measure participants' involvement with and subjective knowledge about the focal products, as well as their online experience levels. Because greater involvement leads to further thoughtfulness regarding the decision outcome and to the more conservative, risk-aversive tendency (Prendergast, Tsang, and Chan 2010), highly involved people may simply be more averse to the riskier disagreement options (Ahluwalia 2002; Wright and Weitz 1977). Conversely, people with high subjective knowledge may exhibit great risk-seeking tendencies for product assessment because they believe that they know enough already about a product and thus tend to increase their confidence and risk-seeking propensity while making their purchase decisions (Duhan et al. 1997; Duncan and Olshavsky 1982; Schmidt and Spreng 1996). Finally, great online experience may foster high risk-seeking tendencies in people towards review ratings by boosting their confidence in online information usage activities; a positive relationship between confidence and risk-seeking tendency has already been established in the literature (Macko

and Tyszka 2009; Shin 2010). Thus, it is important to consider the possibility that the predicted preference of the disagreement over the agreement option may occur simply due to an increased risk-seeking propensity that results from a high level of subjective knowledge/online experience or a low level of involvement, rather than due to a high level of perceived preference heterogeneity. We address these alternative explanations by additionally controlling for such variables in our ANCOVA model.

3.1 Sample and Design

Eighty undergraduate and graduate students participated in the experiment in return for 5,000 won in Korean currency (10 graduate students, 34 women, M_{age} = 22.53). We employed a 2 (ratings variance: disagreement vs. agreement) × 2 (preference heterogeneity: high vs. low) between-subjects design.

3.2 Procedure and Measures

The study was presented as an investigation of online shopping experience. The participants considered a newly published novel or an English listening practice book, depending on their randomly assigned conditions. Both book titles were hypothetical but purported to be for sale at existing online bookstores. The hypothetical bibliographic specifications included the authors' names, publishers' identities, and 10- and 13-

digit ISBN numbers, all of which were provided as aliases. Stock status was fixed as "available" across all conditions. The star rating information subsequently appeared in the form of agreement or disagreement. While the reviewers' ratings consisted of 5 one-star ratings and 5 five-star ratings in the disagreement condition (variance = 4.44, mean = 3), all 10 were three-star ratings in the agreement condition (variance = 0.00, mean = 3; see Appendix A).

After the participants reviewed this information about the titles, they rated their attitude (anchored by "not interesting/interesting," "bad/good," and "not appealing/appealing"; a = .73), and purchase intention (anchored by "very unlikely/very likely" and "very improbable/very probable"; a = .89) regarding the focal title using nine-point bipolar scales in both cases (Chang and Thorson 2004). They also reported their attributions about the review ratings using a single nine-point scale ("To what extent do you feel that the following reasons influenced the reviewers' ratings about the given title?"; "1 = The novel's [or English listening practice book's] quality is the only cause for the review ratings" to "9 = The reviewers' tastes are the only cause for the review ratings"; Mizerski 1978).

Next, they completed manipulation checks of perceived preference heterogeneity using three nine-point scales ("Tastes and preferences are important regarding how people choose a novel

[or English listening practice book]," "A...is a book with which people look for different things," and "Most people want the same thing from... [reverse-coded]"; "1 = not at all" to "9 = very much so"; Price et al. 1989; a = .84). Familiarity was then measured using two nine-point scales ("The above book is new to me [reverse-coded]"; "1 = strongly disagree" to "9 = strongly agree"; Chang and Thorson 2004; and "Are you familiar with the above book title?"; "1 = unfamiliar" to "9 = familiar"; Park and Lessig 1981; a = .76).

Moreover, they indicated their involvement with a book title using three nine-point scales ("I would choose a novel [or an English listening practice book] very carefully," "Deciding which...to buy would be an important decision for me," and "Which...I buy matters to me a lot"; "1 = not at all" to "9 = very much so"; Mittal and Lee 1989; a = .80); reported their levels of online experience using two nine-point scales ("Regardless of your participation in this study, you are experienced with online browsing/shopping"; "1 = not experienced at all" to "9 = very experienced"; Smith, Menon, and Sivarkumar 2005; a = .86); and assessed their subjective knowledge using three nine-point scales ("I know fairly much about a novel [or an English listening practice book]," "I feel very knowledgeable about...," and "Compared to most other people, I know less about...[reverse-coded]"; "1 = not at all" to "9 = very much so"; Flynn and Goldsmith 1999; a = .93).

Finally, the participants provided their demographic information and received a 5,000 won honorarium for their participation.

3.3 Results

3.3.1 Manipulation and Confound Checks

We performed a 2 (ratings variance) × 2 (preference heterogeneity) analysis of variance to check the manipulation of preference heterogeneity. The results confirmed the success of our manipulation; preferences were perceived as more heterogeneous for a novel than for a listening book (M_{novel} = 6.89, $M_{listening}$ = 6.12; $F(1,76)$ = 9.08, $p < .01$; all other effects, Fs $< .75$, ps $> .39$).

Also, we conducted the same three-way ANOVA using familiarity as a dependent variable. The results did not show any significant differences between the two titles (M_{novel} = 8.40, $M_{listening}$ = 8.16; $F(1,76)$ = 1.17, $p > .28$; all other effects, Fs $< .16$, ps $> .69$). Thus, we concluded that a confounding effect of familiarity could not account for the effects we would observe on key dependent variables.

3.3.2 Test of Hypotheses

Consistent with H1, a 2 (ratings variance) × 2 (preference heterogeneity) ANCOVA on attitude using involvement, online experience and subjective knowledge as covariates indicated significant interaction between ratings variance and preference heterogeneity ($F(1,73)$ = 13.75, $p < .01$). No other effects emerged (Fs < 1.65, ps $> .20$). Follow-up contrasts revealed that the participants exposed to the novel reported higher attitudes when the ratings showed disagreement versus agreement (M_{agree} = 3.83, $M_{disagree}$ = 5.23; $F(1,73)$ = 11.91, $p < .01$). Conversely, the participants exposed to the listening book revealed the opposite pattern, albeit at a marginally significant level (M_{agree} = 4.75, $M_{disagree}$ = 4.10; $F(1,73)$ = 2.88, $p < .10$).

Similarly, a parallel ANCOVA on purchase intention indicated significant interaction between ratings variance and preference heterogeneity ($F(1,73)$ = 8.46, $p < .01$). No other effects were significant (Fs < 2.50, ps $> .13$). Planned contrasts revealed that the participants exposed to the novel reported higher levels of purchase intention when the ratings showed disagreement versus agreement (M_{agree} = 2.28, $M_{disagree}$ = 3.73; $F(1,73)$ = 9.33, $p < .01$), but those exposed to the listening book produced opposite, but insignificant, responses (M_{agree} = 2.73, $M_{disagree}$ = 2.20; $F(1,73)$ $< .10$, $p > .30$).

Also, an analogous ANCOVA performed on causal attribution revealed the same interaction between ratings variance and preference heterogeneity, as predicted in H2 ($F(1,73)$ = 5.01, $p < .05$). No other effects were significant (Fs < 1.97, ps $> .16$). Focused contrasts illustrated that when ratings show disagreement versus agreement, the participants exposed to the

novel attributed the review ratings more to reviewers' tastes over product quality (M_{agree} = 5.50, $M_{disagree}$ = 6.40; $F(1,73)$ = 2.69, p = .105), but those exposed to the listening book showed opposite responses (M_{agree} = 5.85, $M_{disagree}$ = 5.10; $F(1,73)$ = 2.21, p = .142), although these differences did not reach conventional levels of significance. Another contrast across ratings variance more explicitly confirmed our prediction in that when the ratings showed disagreement, the participants exposed to the novel attributed the ratings more to the reviewers' tastes to a much greater extent than those exposed to the listening book (M_{novel} = 6.40, $M_{listening}$ = 5.10; $F(1,73)$ = 4.79, $p < .05$), but these differences disappeared when the ratings showed agreement (M_{novel} = 5.50, $M_{listening}$ = 5.85; $F < 1.0$, $p > .40$; see Figure 1).

Overall, these results support our predictions that when preference heterogeneity is perceived as high, disagreement in ratings prompts more favorable product evaluations and enhances causal attribution to reviewers' tastes.[1]

3.3.3 Mediation Analyses

To test the mediating role of causal attribution, we conducted a bootstrapping analysis adopting the PROCESS macro for SPSS (5,000 bootstrap samples; Hayes 2012; Preacher and Hayes 2004). We first performed a bootstrap analysis for attitude, treating ratings variance, preference heterogeneity, their interaction and causal attribution as the independent variables and involvement, online experience, and subjective knowledge as covariates. This analysis showed that the effect of causal attribution on attitude, while controlling for ratings variance and preference heterogeneity was significant (β = .33, t = 4.00, $p < .01$), and the indirect effect of ratings variance and preference heterogeneity via causal attribution was significant (95% confidence interval [CI]: .071, 1.475). Specifically, the conditional indirect effect was significantly positive for the novel (β = .28, 95% CI: .025, .780), but not for the listening book (β = -.25, 95% CI: -.910, .062).

Next, we conducted analogous analyses for purchase intention. The effect of causal attribution on purchase intention, when holding ratings variance and preference heterogeneity constant, was significant (β = .22, t = 2.03, $p < .05$), and the indirect effect of ratings variance and preference heterogeneity via causal attribution was significant (95% CI: .042, .969). Specifically, the conditional indirect effect was

---

1) Parallel results were obtained in a 2 (ratings) × 2 (preference heterogeneity) ANOVA. The analysis of attitude revealed only a significant interaction between ratings variance and preference heterogeneity ($F(1,76)$ = 14.41, $p < .01$; all other effects, Fs < 2.0, $p > .15$). Similarly, the analyses of purchase intention and causal attribution yielded only significant interactions between ratings variance and preference heterogeneity ($F_{purchase\ intention}$ (1,76) = 8.70, $p < .01$; all other effects, Fs < 2.60, $p > .10$; also, $F_{causal\ attribution}$ (1,76) = 5.38, $p < .05$; all other effects, Fs < 1.80, $p > .18$).

<Figure 1> Effect of ratings variance and preference heterogeneity on product evaluation and causal attribution (Study 1).

A. Attitude

B. Purchase intention

C. Causal attribution
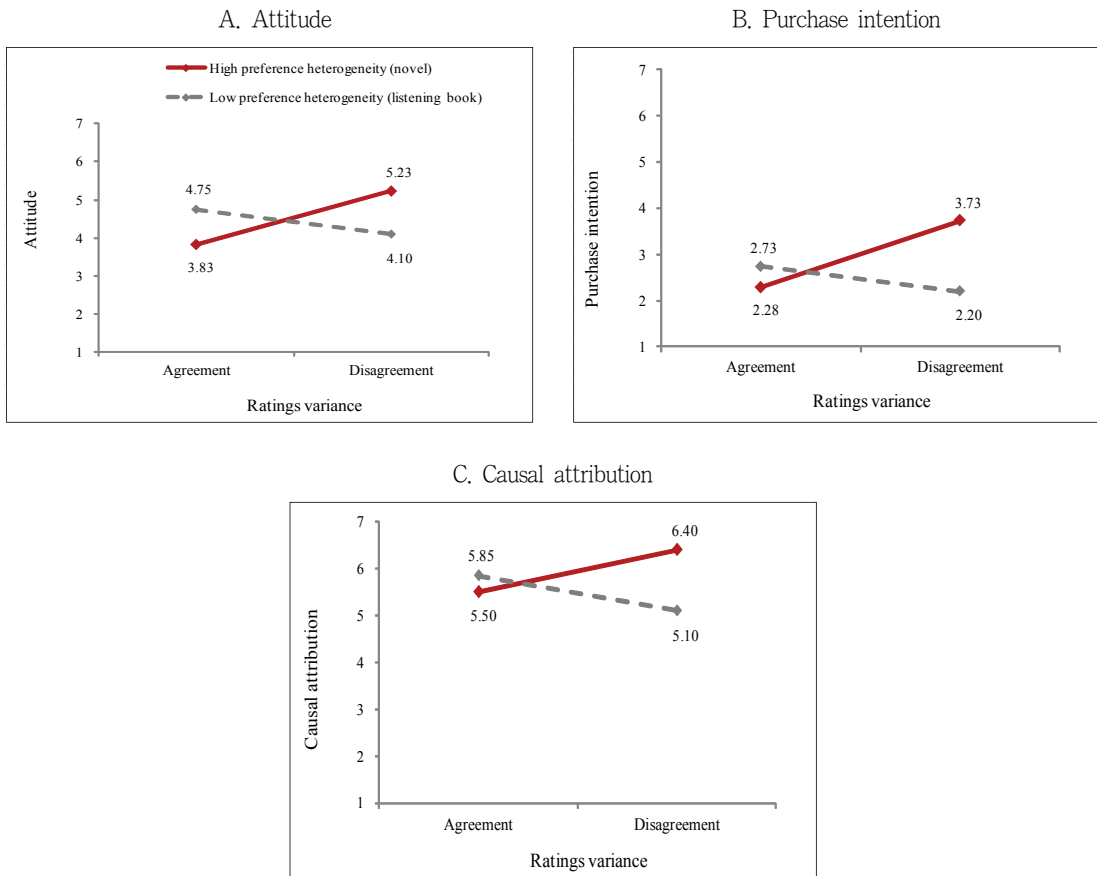
significantly positive for the novel (β = .19, 95% CI: .003, .540), but not for the listening book (β = -.17, 95% CI: -.595, .038).

Overall, these results supported the mediating role of casual attribution in the relationship between ratings variance and preference heterogeneity (H3).

3.4 Discussion

The results of Study 1 provided support for our prediction that disagreement options receive more favorable product evaluations when preference heterogeneity is perceived as high rather than low. More specifically, a bootstrapping analysis confirmed the mediating role of causal attribution, showing that for a high-preference heterogeneity title, disagreement in

ratings is attributed more to reviewers' tastes, which in turn prompts more favorable product evaluations than for a low-preference heterogeneity title.

However, one unexpected finding deserves further attention. For the low-preference heterogeneity title, disagreement in ratings generated similar, rather than lower, product evaluations and causal attribution, compared with agreement in ratings. A possible explanation for this lack of difference is that the type of stimuli titles chosen for a low-preference heterogeneity condition (i.e., the listening book) may have not been sufficiently effective in reducing the perceived preference heterogeneity below a certain threshold to distinguish low- from moderate- to high-preference heterogeneity. Evidence thus far indicates that the choice of treatment stimuli can restrict the range of participants' available behavioral responses (Cesario et al. 2010; Katzko 2006; Zaccaro, Foti, and Kenny 1991). If the researchers' treatment manipulation fails to fully cover the range of possible behavioral responses, participants can respond to the manipulation uniformly below or above a certain threshold (Byrne and Bovair 1997). Thus, even when the manipulation succeeds in dividing participants into two different experimental groups, the range of their responses may still occupy a subset of the potential range, such that only low to moderate responses are obtained if the threshold falls above the manipulation coverage, or only moderate to high

responses are obtained if the threshold falls below it (Elfering, Grebner, and de Tribolet-Hardy 2012).

Our manipulation check results provide some support for the plausibility of this explanation: for a listening book, the value of preference heterogeneity was significantly lower than that for a novel, as evidenced in the manipulation checks ($M_{listening}$ = 6.12, M_{novel} = 6.89, $p <$ .01), but it was still far higher than the scale's midpoint ($M_{listening}$ - midpoint = 6.12 - 5 = 1.12, $t(39)$ = 5.46, $p <$ .01). It thus seems plausible that our manipulation covered only the moderate to high levels of preference heterogeneity, occupying a subset of the range of possible behavioral responses. Our stimuli of a listening book may indeed have aroused a rather moderate level of preference heterogeneity in terms of absolute measures, while arousing a relatively low level of preference heterogeneity in comparison with a novel. We therefore suspect that this insufficiency in the manipulation of preference heterogeneity may account for the lack of statistical significance for a listening book.

In an attempt to overcome this methodological limitation in Study 1, we measure, rather than manipulate, the perceived preference heterogeneity in Study 2. Given that individual responses vary more broadly with self-reported measurement than with experimental manipulation, such a methodological shift would help more fully capture the broad range of individuals'

responses on the focal measure (Eliezer, Major, and Mendes 2010; Elliot and Church 1997). Because a measurement approach permits a more stable assessment of focal effects and the full manifestation of chronic tendencies in the measure of interest, there is likely to be more substantial variation in what participants perceive and report as their own perceived preference heterogeneity (McFerran et al. 2010; Snow 1974). We thus expect that this shift in methodology would help fully cover the broad range of possible responses and increase the generalizability of the findings in a more naturalistic way through the use of a quasi-experimental approach of measuring individual perception of preference heterogeneity (Wirtz and Kimes 2007).

Additionally, in Study 2, we extend our investigations to various levels of average ratings to provide a more comprehensive picture of the interplay among ratings variance, preference heterogeneity, and average ratings. Because consumers tend to engage in casual attribution processing when products meet or exceed their aspiration levels, the causal attribution of ratings variance will be more activated when the average ratings are high enough to reach their aspiration levels (Church et al. 1985; Stephanous and Sage 1987). Therefore, we predict that if products receive high average ratings, the causal attribution of ratings variance will be more activated, and hence its interaction with preference heterogeneity will be more pronounced, as compared to if they receive low average ratings.

Briefly, in Study 2, we measure rather than manipulate the amount of each participant's perceived preference heterogeneity, and extend the experimental designs to various levels of average ratings. We thereby attempt to illustrate how the interplay among ratings variance, preference heterogeneity, and average ratings impacts consumers' responses to online ratings.

Ⅳ. Study 2

Study 2 builds on Study 1 in two important ways. First, we vary the average ratings from two to four stars to test whether the interaction between ratings variance and preference heterogeneity differs as a function of the level of average ratings (H4). By fixing the average ratings at two- or four-star ratings, as well as at three-star ratings, we attempt to make the average ratings higher or lower than a midpoint rating (i.e., three-star ratings; Mudambi and Schuff 2010). This incorporation of various levels of average ratings into the ratings manipulation would help assess the possible three-way interaction between ratings variance, preference heterogeneity, and average ratings.

Second, we measure, rather than manipulate, the degree of perceived preference heterogeneity to tap the wider range of possible responses and to facilitate the full manifestation of chronic dispositions. Specifically, we compare consum-

ers with high and low levels of perceived preference heterogeneity using a hypothetical brand of laptop computer. Because laptop computers incorporate diverse attributes such as the processor speed, amount of RAM, display size, weight, and color, perceived preference heterogeneity can vary depending on what attributes people value most in their decision-making process. While some consumers value measurable attributes such as the processor speed or amount of RAM, others pay more attention to aesthetic features such as the display size, weight, or color (Jiang 2004; Lee and Lee 2011). Thus, preference heterogeneity is likely perceived as rather high when the aesthetic features are emphasized, whereas it is perceived as low when measurable attributes are emphasized (Lee and Lee 2009).

4.1 Sample and Design

Two hundred sixteen undergraduate and graduate students from a large university participated in Study 2 in exchange for partial course credit or monetary compensation (5,000 won; 11 graduate students, 92 women, $M_{age} = 22.48$). The study used a 2 (ratings variance: disagreement vs. agreement) × 3 (average ratings: two-, three-, or four-star ratings) between-subjects experimental design plus a measured preference heterogeneity variable.

4.2 Procedures and Measures

The study was introduced as an investigation of online shopping experience. Participants were presented with hypothetical profiles of a laptop computer, described in terms of its product specifications and reviewers' ratings. Included in the fictitious specifications were the manufacturers' names, model numbers, manufacturing dates, and stock statuses (fixed as "available"). Subsequently presented were the reviewers' ratings, which assumed the form of agreement or disagreement, as in Study 1. Yet, two distinctions were made in Study 2: (1) the ratings volume increased from ten to twenty, and (2) the average ratings were extended to two and four stars, beyond three stars in one step in both directions. Thus, depending on the experimental conditions, participants were presented with the following rating information: in the two-star average rating condition, participants viewed all 20 two-star ratings (agreement; variance = 0.00) or 15 one-star ratings and 5 five-star ratings (disagreement; variance = 3.16). In the three-star average rating condition, they viewed all 20 three-star ratings (agreement; variance = 0.00) or 10 one-star ratings and 10 five-star ratings (disagreement; variance = 4.21). Finally, in the four-star average rating condition, they saw all 20 four-star ratings (agreement; variance = 0.00) or 5 one-star ratings and 15 five-star ratings (disagreement; variance = 3.16; Appendix B).

After exposure to this information, the participants indicated their attitude toward the laptop computer using two nine-point scales anchored by bad/good and not very likable/very likable (Chang and Thorson 2004). They then reported their purchase intention using a single nine-point scale ranging from 1 (not at all likely) to 9 (very likely) (Jain, Buchanan, and Maheswaran 2000). Because attitudes and intentions were highly correlated ($a = .88$, $p < .01$), they were averaged to yield an overall 'product evaluation' measure (Jain et al. 2000).

Next, the participants responded to the causal attribution measure using the same instrument used in Study 1. The perceived preference heterogeneity was also measured with the same instrument used in Study 1 ($a = .72$). Regarding the manipulation check of aspiration levels, a single item was adopted from West and Broniarczyk (1998), which asked participants to provide the minimum standard that they would find acceptable for a laptop computer given the average ratings of the product (1 = horrible laptop computer and 9 = excellent laptop computer).

Also, the involvement ($a = .77$), subjective knowledge ($a = .90$), and online experiment measures ($a = .77$) were all measured using the same instruments from Study 1. The ownership of a laptop computer was then checked with the question, "Do you have your own laptop computer?" (Yes or No). Finally, participants answered demographic questions and were debriefed and dismissed.

4.3 Result

4.3.1 Manipulation and Confound Checks

To check whether the average ratings manipulation influenced participants' aspiration levels, we performed a 2 (ratings variance) × 3 (average ratings) ANOVA, treating the responses to the aspiration levels question as the dependent variable. There was a significant main effect of the average ratings manipulation ($M_{two} = 2.30$, $M_{three} = 3.29$, $M_{four} = 5.63$; $F(2,210) = 113.57$, $p < .01$). A Scheffe post-hoc test revealed that the comparison of the two- versus three-star average ratings conditions ($p < .01$) and the three- versus four-star average ratings conditions ($p < .01$) were both significant. No other effects emerged as significant ($Fs < .70$, $ps > .50$). Hence, we confirmed that the higher average ratings led to higher perceived aspiration levels.

Also, to check whether the incidence of ownership of a laptop computer differed across treatment conditions, we conducted a logit analysis with ratings variance, average ratings, and their interaction as independent variables. No significant treatment effects were found (Wald x^2s $< .20$, $ps > .80$). Thus, we conclude that the differences in the ownership of a laptop computer would not constitute an alternative explanation for the results presented below.

4.3.2 Test of Hypotheses

To test the moderating effect of preference heterogeneity, we performed a series of hierarchical regression analyses. The first model tested the main effects of ratings variance, preference heterogeneity, and average ratings. The second model added the three two-way interaction terms. The third model then added the three-way interaction term. Involvement, subjective knowledge, and online experience were entered into the analyses as covariates. All continuous variables, including the measured predictor (i.e., preference heterogeneity), were centered prior to the analyses to reduce multicollinearity (Aiken and West 1991).

Product evaluations. The results of Model 1 revealed a simple main effect of average ratings in the expected directions (β= 1.42, t = 9.44, $p < .01$). The higher the average ratings, the more positive the product evaluations were. Model 2 then revealed a significant interaction between ratings variance and preference heterogeneity ($\beta = .30$, $t = 2.68$, $p < .01$; H1), but no other interactions ($ps > .40$). Model 3 finally revealed a marginally significant three-way interaction between ratings variance, preference heterogeneity, and average ratings ($\beta = .26$, $t = 1.89$, $p < .06$), which qualified the two-

way interaction (H4a). Adding the three-way interaction to the regression equation marginally significantly increased the amount of explained variance ($R^2 = .012$, $F(1, 205) = 3.59$, $p < .06$; Table 1).

In order to probe the (marginally) significant three-way interaction, we examined the simple interactions between ratings variance and preference heterogeneity at each average-rating condition using the PROCESS macro (Hayes 2012).[2] First tested were the simple two-way interactions at the two-star average rating condition. The analysis showed that there was no significant simple interaction ($\beta = .12$, $SE = .15$, $t = .78$, $p > .40$). Neither the estimated simple slopes of ratings variance for high-preference heterogeneity participants ($\beta = .18$, $SE = .25$, $t = .73$, $p > .45$) nor that for low-preference heterogeneity participants ($\beta = .09$, $SE = .24$, $t = .37$, $p > .70$) was significant, although the directions were as predicted.

However, the simple two-way interactions achieved significance both in the three-star ($\beta = .33$, $SE = .11$, $t = 2.90$, $p < .01$) and four-star ($\beta = .54$, $SE = .17$, $t = 3.21$, $p < .01$) average rating conditions. More specifically, in the three-star average rating condition, high- preference heterogeneity participants reported higher product evaluations when the ratings show disagreement versus

---

2) Regression lines were plotted for one standard deviation above and below the mean for preference heterogeneity (Aiken and West 1991).

agreement ($\beta = .38$, $SE = .17$, $t = 2.16$, $p < .05$), while low-preference heterogeneity participants indicated the opposite pattern ($\beta = .38$, $SE = .18$, $t = 2.08$, $p < .05$). Likewise, in the four-star average rating condition, high-preference heterogeneity participants reported higher product evaluations when they viewed ratings disagreement versus agreement ($\beta = .57$, $SE = .25$, $t = 2.30$, $p < .05$), while low-preference heterogeneity participants reported the opposite pattern ($\beta = .67$, $SE = .27$, $t = 2.49$, $p < .05$; Figure 2).

Causal attribution. A similar procedure to the above was applied to the causal attribution data. Model 1 first revealed a positive effect of average ratings, as was observed in the product evaluation results ($\beta = .30$, $t = 2.06$, $p < .05$). Higher average ratings led to a greater level of attribution to reviewers' tastes. This result is in good agreement with earlier causal attribution studies in that favorable information led to greater attribution to a non-stimulus cause (i.e., reviewers' tastes) over a stimulus cause (i.e., product quality), compared with unfavorable information (Mizerski 1982). Model 2 then revealed a significant interaction between ratings variance and preference heterogeneity ($\beta = .50$, $t = 4.74$, $p < .01$; H2). There were no other interactions ($ps > .75$). Finally, Model 3 confirmed the three-way interaction between ratings variance, preference
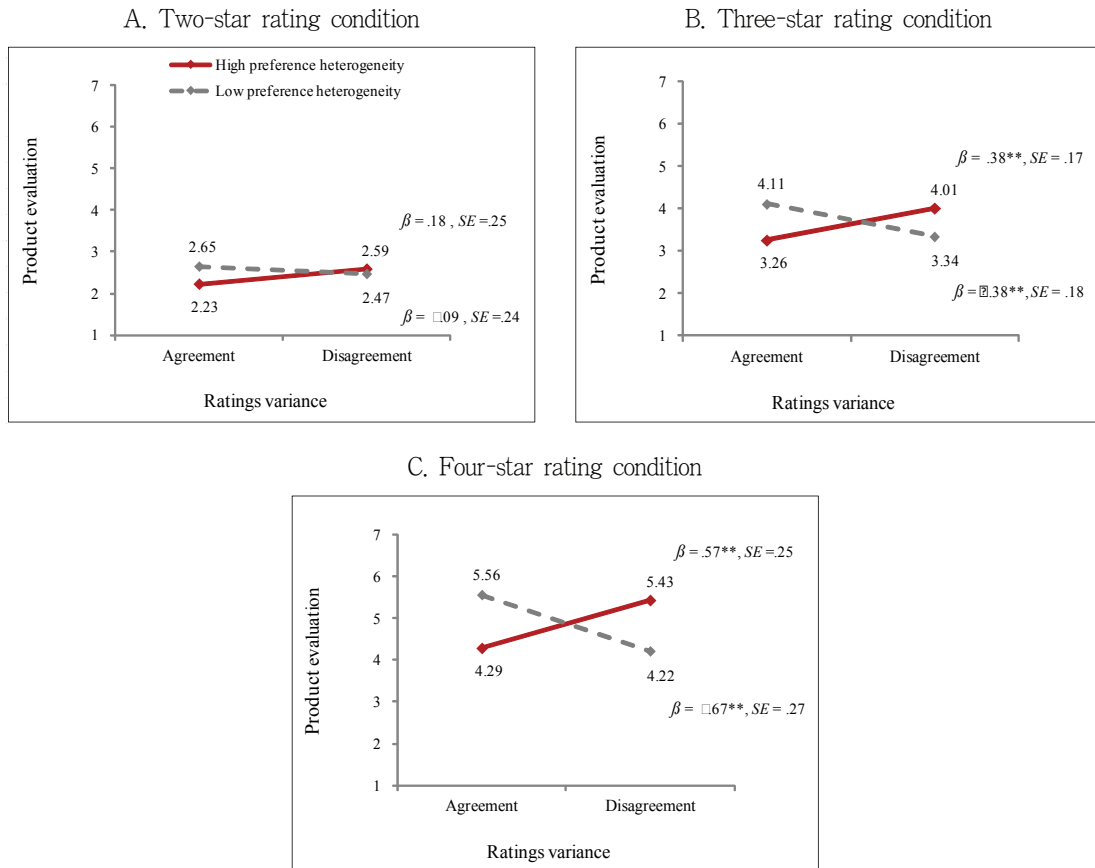
<Table 1> Regression results on product evaluations (Study 2).

Variable	Model 1 β	Model 1 t	Model 2 β	Model 2 t	Model 3 β	Model 3 t
Ratings variance	.03 (.13)	.20	.01(.12)	.08	-.01(.12)	-.03
Preference heterogeneity	.01 (.11)	.11	-.05(.11)	-.47	-.04(.11)	-.34
Average ratings	1.42 (.15)	9.44***	1.42(.15)	9.54***	1.44(.15)	9.73***
Ratings variance × Preference heterogeneity			.30(.11)	2.68***	.33(.11)	2.94***
Ratings variance × Average ratings			-.05(.15)	-.33	-.06(.15)	-.39
Preference heterogeneity × Average ratings			.10(.13)	.78	.03(.14)	.24
Ratings variance × Preference heterogeneity × Average ratings					.26(.14)	1.89*
Involvement	.01(.16)	.08	-.03(.16)	-.20	-.03(.16)	-.20
Subjective Knowledge	-.04(.08)	-.56	-.03(.08)	-.38	-.04(.08)	-.50
Online experience	-.04(.10)	-.35	-.01(.10)	-.01	.02(.10)	.22
F	$F(6, 209) = 15.23$***		$F(3, 206) = 2.50$***		$F(1, 205) = 3.59$***	
$\triangle R^2$	.304		.024		.012	

Notes: *SE* in parentheses
*** $p < .01$. ** $p < .05$. * $p < .10$.

〈Figure 2〉 Effect of ratings variance and preference heterogeneity on product evaluation at various levels of average ratings (Study 2).

Notes: Means estimated +/- 1 SD.
*** $p < .01$. ** $p < .05$.

heterogeneity, and average ratings ($\beta = .35$, $t = 2.79$, $p < .01$; H4b). Adding the three-way interaction to the regression equation significantly increased the amount of explained variance ($R^2 = .032$, $F(1, 205) = 7.75$, $p < .01$; Table 2).

Next, simple interaction analyses were performed to probe the nature of the three-way interaction. In the two-star average ratings condition, there was only marginally significant interaction between ratings variance and preference heterogeneity ($\beta = .25$, $SE = .14$, $t = 1.79$, $p = .075$). Neither the simple slope of ratings variance for high-preference heterogeneity participants ($\beta = .30$, $SE = .23$, $t = 1.33$, $p > .18$) nor that for low- preference heterogeneity participants ($\beta = -.27$, $SE = .22$, $t = -1.19$, $p > .20$) was significant, although the directions of the effects were consistent

with the hypothesis.

However, the same simple interaction analyses yielded significant results in the three-star (β = .53, SE = .10, t = 5.11, p < .01) and four-star (β = .82, SE = .15, t = 5.29, p < .01) average ratings conditions. In the three-star average ratings condition, high-preference heterogeneity participants attributed the review ratings more to the reviewers' tastes when ratings showed disagreement versus agreement (β = .63, SE = .16 , t = 3.92, p < .01), while low-preference heterogeneity participants produced the opposite results (β = -.60, SE = .17, t = -3.55, p < .01). Likewise, in the four-star average rating condition, high-preference heterogeneity participants attributed the

review ratings more to reviewers' tastes when the ratings showed disagreement versus agreement (β = .96, SE = .23, t = 4.18, p < .01), while low-preference heterogeneity participants produced the opposite pattern (β = -.93, SE = .25, t = -3.74, p < .01; Figure 3).

Overall, these results lend support for our hypothesis that the interaction between ratings variance and preference heterogeneity is more pronounced when the average rating is high compared to when it is low (H4).

4.3.3 Mediation Analysis

Similar to Study 1, we tested the mediating role of causal attribution using a bootstrapping

〈Table 2〉 Regression results on casual attribution (Study 2).

Variable	Model 1		Model 2		Model 3	
	β	t	β	t	β	t
Ratings variance	.06 (.12)	.48	.03(.12)	.30	.02(.11)	.14
Preference heterogeneity	.02 (.11)	.14	-.13(.11)	-1.23	-.11(.10)	-1.05
Average ratings	.30 (.14)	2.06**	.30(.14)	2.17**	.33(.14)	2.45**
Ratings variance × Preference heterogeneity			.50(.11)	4.74***	.54(.10)	5.16***
Ratings variance × Average ratings			.01(.14)	.06	-.01(.14)	-.02
Preference heterogeneity × Average ratings			-.04(.12)	-.32	-.13(.12)	-1.07
Ratings variance × Preference heterogeneity × Average ratings					.35(.12)	2.79***
Involvement	-.15(.16)	-.93	-.21(.15)	-1.36	-.21(.15)	-1.38
Subjective Knowledge	.08(.07)	1.02	.12(.07)	1.65	.11(.07)	1.50
Online experience	-.03(.10)	-.27	.02(.09)	.21	.05(.09)	.55
F	$F(6, 209)$ = 1.10		$F(3, 206)$ = 7.77***		$F(1, 205)$ = 7.75***	
$\triangle R^2$	. 031		.098		.032	

Notes: SE in parentheses
*** p < .01. ** p < .05. * p < .10.

<Figure 3> Effect of ratings variance and preference heterogeneity on causal attribution at various levels of average ratings (Study 2).

A. Two-star rating condition

B. Three-star rating condition

C. Four-star rating condition

Notes: Means estimated +/- 1 SD.
*** p < .01. ** p < .05.

procedure (5,000 bootstrap sample; Hayes 2012). This bootstrapping analysis treated product evaluations as a dependent variable; ratings variance, preference heterogeneity, average ratings, their interactions and causal attribution as independent variables; and involvement, online experience and subjective knowledge as covariates. The results showed that the effect of causal attribution remained significant ($\beta = .23$, $t =$

3.03, $p < .01$), but that the three-way interaction was no longer significant ($\beta = .18$, $t = 1.32$, $p > .18$). The indirect effect of the three-way interaction via causal attribution was significant with a 95% bias corrected bootstrap confidence interval of .014 to .188.

More specifically, in the two-star average rating condition, the indirect effect of ratings variance via causal attribution was not sig-

nificant for high- (β = .07, 95% CI: -.027, .220) or low-preference heterogeneity participants (β = -.06, 95% CI: -.213, .023). However, in the three- and four-star average rating conditions, the indirect effects via causal attribution were significantly positive for high-preference heterogeneity participants (β_{three} = .14, 95% CI: .036, .289; β_{four} = .22, 95% CI: .058, .431), while they were negative for low-preference heterogeneity participants (β_{three} = -.14, 95% CI: -.291, -.042; β_{four} = -.21, 95% CI: -.434, -.068). Overall, these results confirm the mediating role of causal attribution in the relationship between ratings variance, preference heterogeneity, and average ratings (H5).

4.4 Discussion

Replicating the findings of Study 1, disagreement options elicited more favorable product evaluations when preference heterogeneity was perceived as high rather than low. Greater causal attribution to reviewers' tastes than to product quality again led to a greater preference for the disagreement option for the high-preference heterogeneity participants. More specifically, this interaction between ratings variance and preference heterogeneity was more pronounced when the average ratings were high rather than low (H4). The mediating role of causal attribution was also confirmed only in the moderate to high level of average ratings conditions, but not in the low average ratings condition (H5).

Additionally, we obtained statistical significance for the low-preference heterogeneity condition in terms of product evaluations and causal attribution, as was shown in the planned comparisons in the three- and four-star average rating conditions. This greater statistical power allows stronger inference regarding the relationship between ratings variance, preference heterogeneity, and average ratings, adequately addressing the lack of statistical significance observed for the low-preference heterogeneity condition in Study 1. The methodological shift from manipulation to measurement appears to contribute to this greater statistical significance by tapping the broader range of participants' responses on the preference heterogeneity measure (Eliezer et al. 2010; McFerran et al. 2010). We thus conclude that the findings obtained here constitute strong evidence for our key prediction regarding the interaction between ratings variance and preference heterogeneity.

V. General Discussion and Implications

5.1 Summary and Implications

Numerous studies have examined the effects of disagreement in ratings or opinions, but the results were not consistent across studies. This

study aims to reconcile these mixed findings by introducing preference heterogeneity as a possible moderator, and causal attribution as a mediator that underlies this moderating effect. We predict, based on such concepts, that product evaluations with disagreement versus agreement in ratings will be more positive when preference heterogeneity is perceived as high rather than low. More specifically, we propose that when preference heterogeneity is perceived as high, disagreement in ratings is attributed more to reviewers' different tastes, and whereby its negative effect on product evaluations is attenuated. We extend this idea to various levels of average ratings, suggesting that the interaction between ratings variance and preference heterogeneity is more pronounced when the average ratings are high, as the likelihood of attaining desired outcomes seems higher and thereby the causal attribution process regarding ratings variance is more activated. Two laboratory studies provided support for the predictions, showing that greater attribution to reviewers' tastes leads to more favorable product evaluations with disagreement in ratings when preference heterogeneity is perceived as high (Study 1) and that the proposed interaction effect between ratings variance and preference heterogeneity is more pronounced when the average rating is high (Study 2).

Our approach has the theoretical merit of recognizing the moderating role of preference heterogeneity when analyzing the effect of

ratings variance. While some results from observational data in studies involving books and game titles confirmed the negative effect of variance (Danescu-Niculescu-Mizil et al. 2009; Zhu and Zhang 2010), other evidence from investigations of fiction books and movies supported the opposite contention (Clement et al. 2007; Martin 2008). To address these contradictions, earlier studies included several moderators, such as aspiration levels (West and Broniarczyk 1998) and product types (Martin 2008; Park and Han 2008). Our work builds on this line of research by proposing and demonstrating that preference heterogeneity moderates the effect of ratings variance on product evaluations and that this moderation varies as a function of the average rating.

Furthermore, we shed light on the mechanism underlying this moderation by linking preference heterogeneity to causal attribution. Research regarding the valence of WOM has examined such a link and reported that taste differences between the WOM sender and receiver increase the attribution of the negativity of WOM to the sender, thereby reducing its negative effect on product evaluations (Liu 2006; Laczniak, DeCarlo, and Ramaswami 2001). We apply this idea to the study of ratings variance, proposing that high preference heterogeneity increases attribution to reviewers' tastes and thereby reduces the negative impact of greater disagreement on product evaluations. Based on this account and on the results supporting it, we at-
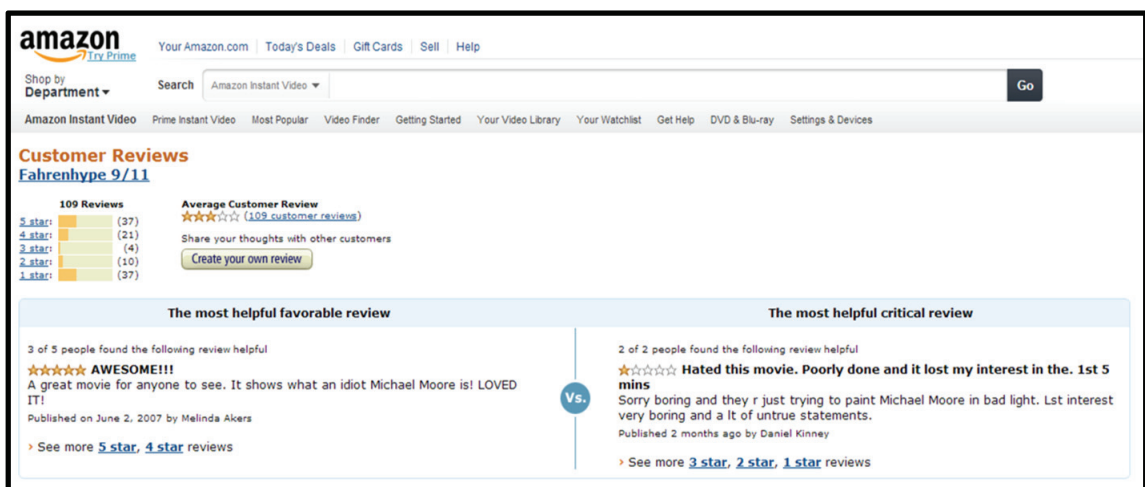
tempt to provide deeper insight into the processes involved in the interaction between ratings variance and preference heterogeneity.

This study also offers important implications for marketers charged with managing eWOM. Faced with dissension among reviewers, consumers cannot entirely avoid the burden of uncertainty surrounding such conflicts. A lack of consensus may leave consumers at a loss in their consumption decision-making processes, increasing their anxiety regarding the choice of a target product (Dugas, Gosselin, and Ladouceur 2001). Our findings suggest that marketers can remove some of these uncertainties by explicitly clarifying that there are many more differences in tastes among reviewers than expected. Because greater attribution to reviewers' tastes help alleviate the psychological burden of unreliable product quality, the uncertainty surrounding the ratings disagreement can be reduced further with this type of attribution.

For example, the review sites of Amazon.com feature the classical "pros and cons" format, directly comparing the most favorable and critical reviews in parallel (See Figure 4). By simultaneously displaying the praise and blame, as well as the respective reasons for either stance, the site more effectively communicates that different preferences may contribute to such controversy. Moreover, Tripadvisor.com, the world's largest travel review site (Stepchenkova and Zhan 2013), offers versatile tools for personalizing the graphical representation of the ratings distribution. The site not only provides the graphs covering the entire sample of reviewers, but also offers filtering tools to personalize the graphs for each customer segment, such as business travelers, families, couples, or friends (See Figure 5). These personalized graphs can help prompt consumers to visualize

〈Figure 4〉 Screenshot of Amazon.com review site (accessed at August 12, 2013).

the ratings distribution per each segment more concretely and facilitate the causal inference of what specific segment hold deviant preferences from those of the majority. Therefore, if given the chance to customize the review representation according to their own criteria, such as reviewers' levels of similarity in terms of the trip purpose, the activities engaged in during a trip, or demographics (Gretzel, Yoo, and Purifoy 2007), consumers would more readily infer that ratings disagreement simply reflects the various preferences among different subgroups, rather than the unreliability of the service quality.

5.2 Limitations and Future research

A few caveats should be noted when interpreting the results of this study. First, our stimuli products did not cover a broad spectrum of price; this was confined to the mid-level price category. For high-priced items such as dental implants or premium jewelry, however, consumers are bound to experience much higher levels of financial risk as compared to low- to middle-value products such as book titles or laptop computers. The greater the financial risk, the greater the risk-averseness of consumers toward the target product (Choi and Ruszczyński 2011). In fact, if preference heterogeneity is perceived as low, as in dental implants, such greater risk averseness does not raise critical concerns because it would not alter our original proposition that for low-prefer-

ence heterogeneity products, the riskier disagreement option is less desirable compared to the safer agreement option. However, when preference heterogeneity is perceived as high, as in premium jewelry, greater risk-averseness may reverse the proposed preference of disagreement over agreement such that consumers rigidly stick to the safer agreement option despite the possibility that different preferences, not a low reliability of the product quality, contribute to the dissension associated with the other option. Another attempt to employ higher-priced items as focal products should help disentangle this effect of financial risk on consumers' responses to disagreement in ratings.

We also note that we only used student samples across the two studies and thereby ensured internal validity but somewhat compromised external validity. With a homogeneous sample set, we could obtain more significant differences by treatment effects, more effectively controlling for possible extraneous variables and substantiating the proposed causal relationship (Wang and Yang 2008). Given that this study primarily focuses on the underlying psychological processes involved in online customer ratings, this emphasis on internal over external validity is warranted (Matula, Pratt, and Sautter 1995). Nonetheless, an attempt to increase external validity, such as through an examination of another age group, can highlight another pathway to extend the results of this study. For example, tolerance levels for di-

versity or individual differences are known to decrease as one ages (Spaeth, 1969). Therefore, confronted with disagreement in ratings, older people may yield more conservative responses and undermine product evaluations overall due to their failure to accept the possibility that various preferences govern a market. In this regard, the inclusion of non-student samples, specifically those of the elderly population, can help gain deeper insight into the roles that individual characteristics play in online ratings dynamics.

References

Ahluwalia, Rohini (2002), "How Prevalent Is the Negativity Effect in Consumer Environments?" *Journal of Consumer Research*, 29(2), 270-279.

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage.

Arsego, Martha Brito (2009), "Reading in Jane Austen's Novels," *Revista da Graduação*, 2(2).

Bae, Jungho, Bum Jun Shim, and Byung-Do Kim (2010), "Simultaneous Effect Between eWOM and Revenues: Korea Movie Industry," *Asia Marketing Journal*, 12(2), 1-25.

Byrne, Michael D. and Susan Bovair (1997), "A Working Memory Model of a Common Procedural Error," *Cognitive Science*, 21(1), 31-61.

Calder, Bobby J. and Robert E. Burnkrant (1977), "Interpersonal Influence on Consumer Behavior: An Attribution Theory Approach," *Journal of Consumer Research*, 4(1), 29-38.

Cesario, Joseph, Jason E. Plaks, Nao Hagiwara, Carlos David Navarrete, and E. Tory Higgins (2010), "The Ecology of Automaticity: How Situational Contingencies Shape Action Semantics and Social Behavior," *Psychological Science*, 21(9), 1311-1317.

Chaiken, Shelly and Charles Stangor (1987), "Attitudes and Attitude Change," *Annual Review of Psychology*, 38(1), 575-630.

Chang, Yuhmiin and Esther Thorson (2004), "Television and Web Advertising Synergies," *Journal of Advertising*, 33(2), 75-84.

Chatterjee, Patrali (2001), "Online Reviews: Do Consumers Use Them?" in *Advances in Consumer Research*, Vol. 28, Mary C. Gilly and Joan Myers-Levy, eds. Provo, UT: Association for Consumer Research, 129-134.

Cheng, Xiufang, and Meihua Zhou (2010), "Study on Effect of eWom: A Literature Review and Suggestions for Future Research," in *Proceedings of the 4th Inter-*

national Conference on Management and Service Science, MASS‑10. Wuhan, China: IEEE, 1-4.

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

Choi, Sungyong and Andrzej Ruszczyński (2011), "A Multi-Product Risk-Averse Newsvendor with Exponential Utility Function," *European Journal of Operational Research*, 214(1), 78-84.

Church, Nancy J., Michel Laroche, and Jerry A. Rosenblatt (1985), "Consumer Brand Categorization for Durables with Limited Problem Solving: An Empirical Test and Proposed Extension of the Brisoux-Laroche Model," *Journal of Economic Psychology*, 6(3), 231-253.

Clement, Michel, Dennis Proppe, and Armin Rott (2007), "Do Critics Make Bestsellers? Opinion Leaders and the Success of Books," *Journal of Media Economics*, 20(2), 77-105.

Cui, Geng, Hon-Kwong Lui, and Xiaoning Guo (2012), "The Effect of Online Consumer Reviews on New Product Sales," *International Journal of Electronic Commerce*, 17(1), 39-58.

Danescu-Niculescu-Mizil, Cristian, Gueorgi Kossinets, Jon Kleinberg, and Lillian Lee (2009), "How Opinions Are Received by Online Communities: A Case Study On Amazon. com Helpfulness Votes," in *Proceedings of the 18th international conference on World wide Web, ACM 2009*. Madrid, Spain: International World Wide Web Conference Committee, 141-150.

Deater-Deckard, Kirby, Kenneth A. Dodge, and Emma Sorbring (2005), "Cultural Differences in the Effects of Physical Punishment," in *Ethnicity and Causal Mechanisms*, M. Rutter and M. Tienda, eds. New York, NY: Cambridge University Press, 204-226.

Dellarocas, Chrysanthos, Neveen F. Awad, and Xiaoquan Michael Zhang (2004), "Exploring the Value of Online Reviews to Organizations: Implications for Revenue Forecasting and Planning," in *Proceedings of the 25th International Conference on Information Systems, ACM 2004*. New York, NY: ACM Press, 379-386.

_____, _____, Xiaoquan Michael Zhang, and Neveen F. Awad (2007), "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing*, 21(4), 23-45.

Duan, Wenjing, Bin Gu, and Andrew B. Whinston (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.

Dugas, Michel J., Patrick Gosselin, and Robert Ladouceur (2001), "Intolerance of Uncer-

tainty and Worry: Investigating Specificity in a Nonclinical Sample," *Cognitive Therapy and Research*, 25(5), 551-558.

Duhan, Dale F., Scott D. Johnson, James B. Wilcox, and Gilbert D. Harrell (1997), "Influences on Consumer Use of Word-of-Mouth Recommendation Sources," *Journal of the Academy of Marketing Science*, 25(4), 283-295.

Duncan, Calvin P. and Richard W. Olshavsky (1982), "External Search: The Role of Consumer Beliefs," *Journal of Marketing Research*, 19(1), 32-43.

Elfering, Achim, Simone Grebner, and Fanny de Tribolet-Hardy (2012), "The Long Arm of Time Pressure at Work: Cognitive Failure and Commuting Near-Accidents," *European Journal of Work and Organizational Psychology*, (ahead-of-print), 1-13.

Eliezer, Dina, Brenda Major, and Wendy Berry Mendes (2010), "The Costs of Caring: Gender Identification Increases Threat Following Exposure to Sexism," *Journal of Experimental Social Psychology*, 46(1), 159-165.

Elliot, Andrew J. and Marcy A. Church (1997), "A Hierarchical Model of Approach and Avoidance Achievement Motivation," *Journal of Personality and Social Psychology*, 72 (1), 218-232.

Erickson, Bonnie H. (1996), "Culture, Class, and Connections," *American Journal of Sociology*, 217-251.

ETnews (2005), "An Investigation of Online Review Usage Behavior," *ETnews.COM*, (November 2), (accessed February 3, 2013), [available at http://news.naver.com /main /read.nhn?mode=LSD&mid=sec&sid1= 105&oid=030&aid=0000122441].

Feick, Lawrence and Robin A. Higie (1992), "The Effects of Preference Heterogeneity and Source Characteristics on Ad Processing and Judgments about Endorsers," *Journal of Advertising*, 21(2), 9-24.

Flynn, Leisa Reinecke and Ronald E. Goldsmith (1999), "A Short, Reliable Measure of Subjective Knowledge," *Journal of Business Research*, 46(1), 57-66.

Folkman, Susan and Judith Tedlie Moskowitz (2004), "Coping: Pitfalls and Promise," *Annual Review of Psychology*, 55, 745-774.

Freeland-Graves, Jeanne and Susan Nitzke (2002), "Position of the American Dietetic Association: Total Diet Approach to Communicating Food and Nutrition Information," *Journal of the American Dietetic Association*, 102(1), 100-108.

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

_____, _____ and José C. Silva (2012), "Sequential and Temporal Dynamics of Online Opinion," *Marketing Science*, 31 (3), 448-473.

Gretzel, Ulrike, Kyung Hyan Yoo, and Melanie Purifoy (2007), "Online Travel Review Study: Role and Impact of Online Travel Reviews," Laboratory for Intelligent Systems in Tourism, Texas University.

Hann, Il-Horn, Kai-Lung Hui, Sang-Yong Tom Lee, and Ivan P. L. Png (2007), "Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach," *Journal of Management Information Systems*, 24(2), 13-42.

Hayes, Andrew F. (2012), *My Macros and Code for SPSS and SAS*, (accessed December 10, 2012), [available at http://www.afhayes.com/spss-sas-and-mplus-macros-and-code.html].

He, Stephen Xihao (2012), "Consumer Judgment and Forecasting Using Online Word-of-Mouth," Ph. D Thesis, Geogia Institute of Technology, Atlanta, Georgia, (accessed February 9, 2013), [available at https://smartech.gatech.edu/bitstream/handle/1853/44866/he_stephenxihao_201208_phd.pdf].

Hennig-Thurau, Thorsten, Kevin P. Gwinner, Gianfranco Walsh, and Dwayne D. Gremler, (2004), "Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet?" *Journal of Interactive Marketing*, 18(1), 38-52.

Herr, Paul M., Frank R. Kardes, and John Kim (1991), "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective," *Journal of Consumer Research*, 17(4), 454-462.

Holbrook, Morris B. (1999), "Popular Appeal Versus Expert Judgments of Motion Pictures," *Journal of Consumer Research*, 26(2), 144-155.

Huang, Peng, Nicholas H. Lurie, and Sabyasachi Mitra (2009), "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods," *Journal of Marketing*, 73(2), 55-69.

Iglesias, Victor (2009), "The Attribution of Service Failures: Effects on Consumer Satisfaction," *Service Industries Journal*, 29(2), 127-141.

Jain, Shailendra Pratap, Bruce Buchanan and Durairaj Maheswaran (2000), "Comparative Versus Noncomparative Advertising: The Moderating Impact of Prepurchase Attribute Verifiability," *Journal of Consumer Psychology*, 9(4), 201-211.

Jiang, Pingjun (2004), "The Role of Brand Name in Customization Decisions: A Search vs Experience Perspective," *Journal of Product & Brand Management*, 13(2), 73-83.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica: Journal of the Econometric Society*, 47(2), 263-292.

Karniouchina, Ekaterina V. (2011), "Impact of Star and Movie Buzz on Motion Picture

Distribution and Box Office Revenue," *International Journal of Research in Marketing*, 28(1), 62-74.

Katzko, Michael W. (2006), "A Study of the Logic of Empirical Arguments in Psychological Research: 'The Automaticity of Social Behavior' as a Case Study," *Review of General Psychology*, 10(3), 210-228.

Keller, Ed (2007), "Unleashing the Power of Word of Mouth: Creating Brand Advocacy to Drive Growth," *Journal of Advertising Research*, 47(4), 448-452.

Kirchler, Erich, Stephan Muehlbacher, Erik Hoelzl, and Paul Webley (2009), "Effort and Aspirations in Tax Evasion: Experimental Evidence," *Applied Psychology: An International Review*, 58(3), 488-507.

Klein, Noreen M. and Stewart W. Bither (1987), "An Investigation of Utility-Directed Cutoff Selection," *Journal of Consumer Research*, 14(2), 240-256.

Laczniak, Russell N., Thomas E. DeCarlo, and Sridhar N. Ramaswami (2001), "Consumers' Responses to Negative Word-of-Mouth Communication: An Attribution Theory Perspective," *Journal of Consumer Psychology*, 11(1), 57-73.

Lee, Byung-Kwan and Wei-Na Lee (2011), "The Impact of Product Knowledge on Consumer Product Memory and Evaluation in the Competitive Ad Context: The Item-Specific-Relational Perspective," *Psychology & Marketing*, 28(4), 360-387.

Lee, Jin Kyun, and Wei-Na Lee (2009), "Country-of-Origin Effects on Consumer Product Evaluation and Purchase Intention: The Role of Objective Versus Subjective Knowledge," *Journal of International Consumer Marketing*, 21(2), 137-151.

Li, Charlene and Josh Bernoff (2008), *Groundswell: Winning in a World Transformed by Social Technologies*. Boston, MA: Harvard Business Press.

Lin, Tom M.Y., Pin Luarn, and Yun Kuei Huang (2005), "Effect of Internet Book Reviews on Purchase Intention: A Focus Group Study," *Journal of Academic Librarianship*, 31(5), 461-468.

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

Maheswaran, Durairaj and Shelly Chaiken (1991), "Promoting Systematic Processing in Low-Motivation Settings: Effect of Incongruent Information on Processing and Judgment," *Journal of Personality and Social Psychology*, 61(1), 13-25.

Macko, Anna and Tadeusz Tyszka (2009), "Entrepreneurship and Risk Taking," *Applied Psychology: An International Review*, 58(3), 469-487.

Martin, Jolie Mae (2008), "Seeing the Forest for the Trees: Information Aggregation in Online Decision-Making," doctoral dissertation, Harvard University, Cambridge, MA.

Martin, Leonard L. and Blossom Davies (1998), "Beyond Hedonism and Associationism: A Configural View of the Role of Affect in Evaluation, Processing, and Self-Regulation," *Motivation and Emotion*, 22(1), 33-51.

Matula, Thomas, Eric R. Pratt, and Elise Truly Sautter (1995), "The Use of Non-Traditional Students as Indicators of the Generalizability of Results in Marketing Research Studies Using Student Samples," in *Proceedings of the Western Marketing Educators' Association Conference, Interactive Marketing: Tuned in, On Line, and Going Places*, San Diego, California: Western Marketing Educators' Association, 66-69.

McFerran, Brent, Darren W. Dahl, Gerald J. Gorn, and Heather Honea (2010), "Motivational Determinants of Transportation into Marketing Narratives," *Journal of Consumer Psychology*, 20(3), 306-316.

Mediatoday (2006), "Eighty Percent of Online Population Read Online News Comments," *Mediatoday*, (September 22), (accessed February 3, 2013), [available at http://www.mediatoday.co.kr/news/articleView.html?idxno=50247].

Meyers-Levy, Joan and Brian Sternthal (1991), "Gender Differences in the Use of Message Cues and Judgments," *Journal of Marketing Research*, 28(1), 84-96.

Mittal, Banwari and Myung-Soo Lee (1989), "A Causal Model of Consumer Involvement," *Journal of Economic Psychology*, 10(3), 363-389.

Mizerski, Richard W. (1978), "Causal Complexity: A Measure of Consumer Causal Attribution," *Journal of Marketing Research*, 15(2), 220-228.

_____, _____ (1982), "An Attribution Explanation of the Disproportionate Influence of Unfavorable Information," *Journal of Consumer Research*, 9(3), 301-310.

Moe, Wendy W. and David A. Schweidel (2012), "Online Product Opinions: Incidence, Evaluation, and Evolution," *Marketing Science*, 31(3), 372-386.

____, _____ and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48(3), 444-456.

Mudambi, Susan M. and David Schuff (2010), "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com.," *MIS Quarterly*, 34(1), 185-200.

Oliver, R.L. (1997), *Satisfaction: A behavioral perspective on the consumer*. Boston, MA: McGraw-Hill.

Park, Cheol and Thae Min Lee (2009), "Information Direction, Website Reputation and eWOM Effect: A Moderating Role of Product Type," *Journal of Business Research*, 62(1), 61-67.

Park, C. Whan and V. Parker Lessig (1981), "Familiarity and Its Impact on Consumer Decision Biases and Heuristics," *Journal of Consumer Research*, 8(2), 223-231.

Park, Do-Hyung and Ingoo Han (2008), "Integrating Conflicting Reviews: Attributional Hypotheses of Consumer Response to Information Uncertainty Depending on Prior Brand Attitude," in *Proceedings of the 41st Hawaii International Conference on System Sciences*, Waikoloa, Hawaii: IEEE Computer Society, 1-10.

Patzelt, Holger and Dean A. Shepherd (2008), "The Decision to Persist with Underperforming Alliances: The Role of Trust and Control," *Journal of Management Studies*, 45(7), 1217-1243.

Preacher, Kristopher J. and Andrew F. Hayes (2004), "SPSS and SAS Procedures for Estimating Indirect Effects in Simple Mediation Models," *Behavior Research Methods*, 36(4), 717-731.

Prendergast, Gerard P., Alex S.L. Tsang, and Cherry N.W. Chan (2010), "The Interactive Influence of Country of Origin of Brand and Product Involvement on Purchase Intention," *Journal of Consumer Marketing*, 27(2), 180-188.

Price, Linda L., Lawrence F. Feick, and Robin A. Higie (1989), "Preference Heterogeneity and Coorientation as Determinants of Perceived Informational Influence," *Journal of Business Research*, 19(3), 227-242.

Ryu, Gangseog, Jongchul Park, and Lawrence Feick (2006), "The Role of Product Type and Country-of-Origin in Decisions About Choice of Endorser Ethnicity in Advertising," *Psychology & Marketing*, 23(6), 487-513.

Scheier, Michael F. and Charles S. Carver (1987), "Dispositional Optimism and Physical Well-Being: The Influence of Generalized Outcome Expectancies on Health," *Journal of Personality*, 55(2), 169-220.

Schindler, Robert M., Michael Berbaum, and Donna R. Weinzimer (1987), "How an Attention-Getting Device Can Affect Quick Choice among Similar Alternatives," in *Advances in Consumer Research*, Vol. 14, Melanie Wallendorf and Paul Anderson, eds. Provo, UT: Association for Consumer Research, 505-509.

Schmidt, Jeffrey B. and Richard A. Spreng (1996), "A Proposed Model of External Consumer Information Search," *Journal of the Academy of Marketing Science*, 24(3), 246-256.

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

Senko, Corwin and Chris S. Hulleman (2013), "The Role of Goal Attainment Expectancies in Achievement Goal Pursuit," *Journal of Educational Psychology*, 105(2), 504-521.

Sideridis, Georgios D. and Avi Kaplan (2011), "Achievement Goals and Persistence Across Tasks: The Roles of Failure and Success," *Journal of Experimental Education*,
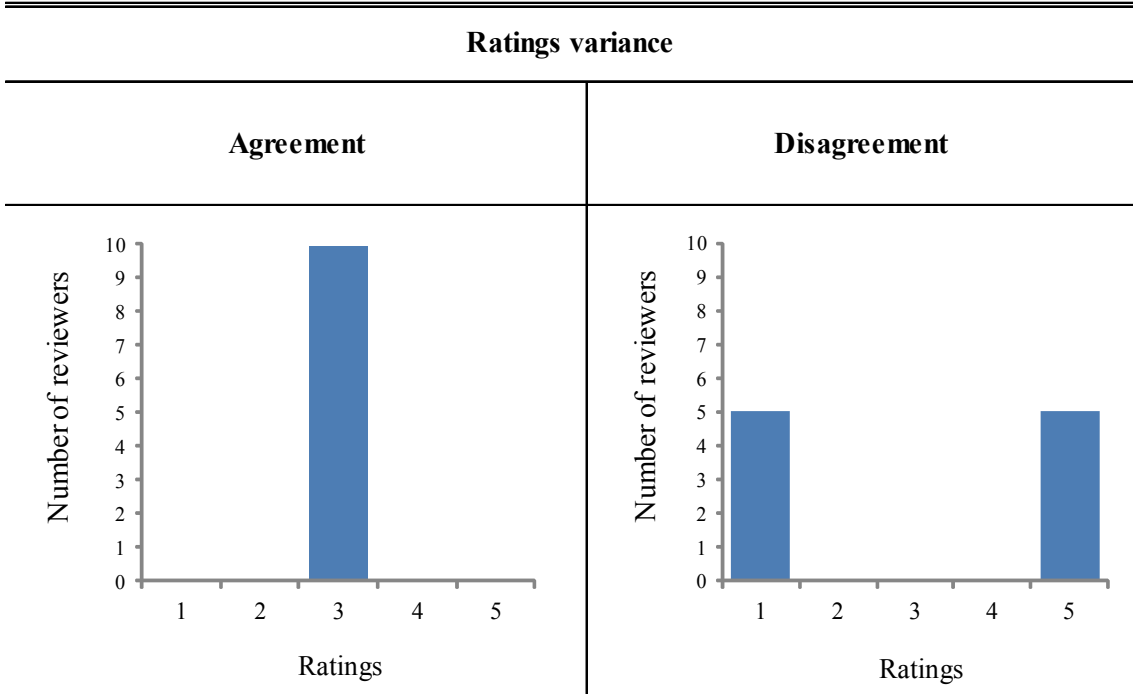
79(4), 429-451.

Shin, Dong-Hee (2010), "Analysis of Online Social Networks: A Cross-National Study," *Online Information Review*, 34(3), 473-495.

Smith, Donnavieve, Satya Menon, and K. Sivakumar (2005), "Online Peer and Editorial Recommendations, Trust, and Choice in Virtual Markets," *Journal of Interactive Marketing*, 19(3), 15-37.

Snow, Richard E. (1974), "Representative and Quasi-Representative Designs for Research on Teaching," *Review of Educational Research*, 44(3), 265-291.

Spaeth, Joe L. (1969), "Public Reactions to College Student Protests," *Sociology of Education*, 42(2), 199-206.

Stepchenkova, Svetlana and Fangzi Zhan (2013), "Visual Destination Images of Peru: Comparative Content Analysis of DMO and User-Generated Photography," *Tourism Management*, 36(June), 590-601.

Stephanou, Harry E. and Andrew P. Sage (1987), "Perspectives on Imperfect Information Processing," *IEEE Transactions on Systems, Man and Cybernetics*, 17(5), 780-798.

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

Taylor, Shelley E. and Susan T. Fiske (1978), "Salience, Attention, and Attribution: Top of the Head Phenomena," in *Advances in Experimental Social Psychology*, Vol. 11,

L. Berkowitz, ed. San Diego, CA: Academic Press, 249-288.

Thomas, Brian Jasen (2007), *Food Deserts? Purchasing Patterns and Perceptions of Food Insecure and Secure Households in a Midwest Urban Neighborhood*. Ann Arbor, MI: ProQuest.

Wang, Xuehua and Zhilin Yang (2008), "A Meta-Analysis of Effect Sizes in International Marketing Experiments," *International Marketing Review*, 25(3), 276-291.

West, Patricia M. and Susan M. Broniarczyk (1998), "Integrating Multiple Opinions: The Role of Aspiration Level on Consumer Response to Critic Consensus," *Journal of Consumer Research*, 25(1), 38-51.

Wicker, Bruno, Christian Keysers, Jane Plailly, Jean-Pierre Royet, Vittorio Gallese, and Giacomo Rizzolatti (2003), "Both of Us Disgusted in My Insula: The Common Neural Basis of Seeing and Feeling Disgust," *Neuron*, 40(3), 655-664.

Wirtz, Jochen and Sheryl E. Kimes (2007), "The Moderating Role of Familiarity in Fairness Perceptions of Revenue Management Pricing," *Journal of Service Research*, 9(3), 229-240.

Wright, Alice A. and John G. Lynch Jr. (1995), "Communication Effects of Advertising Versus Direct Experience," *Journal of Consumer Research*, 21(4), 708-718.

Wright, Peter and Barton Weitz (1977), "Time Horizon Effects on Product Evaluation

Strategies," *Journal of Marketing Research*, 14(4), 429-443.

Yap, Kenneth B., Budi Soetarto, and Jillian C. Sweeney (2013), "The Relationship Between Electronic Word-of-Mouth Motivations and Message Characteristics: The Sender's Perspective," *Australasian Marketing Journal*, 21(1), 66-74.

Zaccaro, Stephen J., Roseanne J. Foti, and David A. Kenny (1991), "Self-Monitoring and Trait-Based Variance in Leadership: An Investigation of Leader Flexibility Across Multiple Group Situations," *Journal of Applied Psychology*, 76(2), 308-315.

Zhu, Feng and Xiaoquan Zhang (2010), "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, 74(2), 133-148.

Appendix A

Manipulation of ratings variance (Study 1).

Ratings variance	
Agreement	**Disagreement**

Appendix B

Manipulation of ratings variance (Study 2).

	Ratings variance	
Average rating	**Agreement**	**Disagreement**
Two-star rating		
Three-star rating		
Four-star rating		