DREADING AND RANTING: THE DISTINCT EFFECTS OF ANXIETY AND ANGER IN ONLINE SELLER REVIEWS

Completed Research Paper

Dezhi Yin
College of Management
Georgia Institute of Technology
Atlanta, GA 30308
dezhi.yin@mgt.gatech.edu

Samuel Bond
College of Management
Georgia Institute of Technology
Atlanta, GA 30308
sam.bond@mgt.gatech.edu

Han Zhang
College of Management
Georgia Institute of Technology
Atlanta, GA 30308
han.zhang@mgt.gatech.edu

Abstract

This paper explores effects of the emotions embedded in a seller review on its perceived helpfulness. Drawing on frameworks from the emotion and cognitive processing literatures, the authors propose that although emotional review content is subject to a well-known negativity bias, the effects of discrete emotions will vary, and that one source of this variance is perceptions of reviewers’ cognitive effort. We focused on the roles of two distinct, negative emotions common to seller reviews: anxiety and anger. In Study 1, actual seller reviews from Yahoo Shopping websites were collected to determine the effects of anxiety and anger on review helpfulness. In Study 2, an experiment was utilized to identify and explain the differential impact of anxiety and anger in terms of perceived reviewer effort. Our findings demonstrate the importance of examining discrete emotions in online word-of-mouth, and they also carry important practical implications for consumers and online retailers.

Keywords: discrete emotions, seller reviews, user-generated content, review helpfulness, online word-of-mouth, electronic commerce, consumer decision-making
Introduction

Online reviews play an increasingly important role in the popularity and success of electronic commerce. In a manner similar to other forms of online word-of-mouth, online reviews help reduce uncertainty surrounding the shopping experience and inform future consumers (Dellarocas 2003). Specifically, the availability of opinions from previous customers can boost buyer trust (Ba and Pavlou 2002) and increase sales (Chevalier and Mayzlin 2006; Forman et al. 2008). For purposes of this paper, 'online reviews' refer to peer-generated evaluations posted on company or third party websites (Mudambi and Schuff 2010). Furthermore, although online reviews can be targeted towards either products or sellers, we focus on the latter which have received surprisingly limited scholarly attention (e.g., Ba and Pavlou 2002; Pavlou and Dimoka 2006; Qu et al. 2008).

The present research concerns determinants of the perceived helpfulness of seller reviews. Online retailers and third party sites often provide voting mechanisms to identify those reviews that are most useful for assisting consumers in their purchase decisions. In the domain of product reviews, many online retailer interfaces (such as that used by Amazon) sort user reviews according to their helpfulness ratings by default. In the domain of seller reviews, many e-commerce platforms have begun to offer this functionality as well (e.g., Yahoo! Shopping Merchant Reviews, ResellerRatings, Google Checkout Reviews). For instance, after each review presented in Yahoo! Shopping, the interface asks “Was this review helpful?” and displays the rating assigned by prior readers (e.g., “4 out of 7 found this review helpful”). Following Mudambi and Schuff’s (2010) definition of a helpful product review, we define a helpful seller review as a peer-generated seller evaluation that facilitates the consumer’s purchase decision process.

Online reviews are unique from other forms of word-of-mouth in that a vast number of reviews are often available, and their authors are usually unknown to the reader. There can be hundreds of thousands of reviews for a single seller. Such widespread availability could provide more information to customers, but also may create problems such as information overload (Jones et al. 2004). The sheer amount of reviews with varying content and quality makes it impossible for customers to evaluate and comprehend all the information before making a purchase decision (Liu et al. 2008). Therefore, voting systems can assist prospective consumers seeking to make an informed purchase decision with minimum effort. In many cases, consumers may only require a small set of helpful reviews, which tend to be more persuasive in shaping readers’ attitude and purchase behaviors (Chen et al. 2008). The possibility of sorting reviews based on helpfulness enables potential consumers to shorten their information search, evaluate alternatives more efficiently, and make better purchase decisions (Cao et al. 2011; Mudambi and Schuff 2010). By providing diagnostic information across all stages of the decision-making process, the mechanism adds value to prospective customers. In one prominent example, it is estimated that Amazon.com added $2.7 billion to annual revenue by asking the simple question “Was this review helpful to you?” (Spool 2009).

A better understanding of the “helpfulness” of seller reviews is likely to benefit online retailers and third-party review providers. Helpful reviews tend to be read more frequently and weighted more heavily by prospective customers. Therefore, if the most helpful reviews of an online retailer are mostly positive, the retailer can expect benefits in terms of reputation, trust, and sales. In contrast, if the most helpful reviews are largely critical, the retailer is likely to suffer. Online retailers can rely on voting mechanisms to identify helpful reviews, but the accumulation of votes takes time (Zhang and Tran 2010). Retailers would prefer to identify more helpful reviews (especially negative ones) early on, even before votes have accumulated, in order to receive the feedback benefits as soon as possible and react accordingly. In addition, review providers themselves stand to gain by providing high quality reviews that bring potential value to both customers and retailers. Customers are more likely to visit and use review websites if doing so can help to mitigate uncertainty/risk, reduce search effort, and enable them to make better decisions more efficiently (Dabholkar 2006). Websites that provide more helpful information than competitors will gain a strategic advantage in attracting consumers’ attention and increasing their time spent reading the reviews (i.e., site “stickiness”) (Connors et al. 2011). Understanding the determinants of review helpfulness is essential for developing writing guidelines to encourage more useful seller reviews. Customers who wish to leave reviews for the benefit of future shoppers may have little awareness regarding what constitutes a helpful...
review. Moreover, the added value of helpful reviews makes them profitable when sold to retailers who thrive to attract new customers and increase sales (Mudambi and Schuff 2010).

Scholars investigating review helpfulness have focused primarily on a number of determinants that are easily observable, such as ratings and reviewer characteristics (Chevalier and Mayzlin 2006; Forman et al. 2008; Mudambi and Schuff 2010), and a few studies have investigated the content and substance of reviews themselves (Cao et al. 2011; Pavlou and Dimoka 2006). A common finding across these studies is a negativity bias whereby negative reviews tend to be more influential than positive ones. However, the variables used in all of these studies are non-emotional (e.g., numerical ratings, text comment meanings). In particular, prior work has tended to regard ‘negativity’ as a global construct, and no research has examined the distinctive roles of various negative emotions contained in negative reviews.

Extending earlier work, we suggest that the emotions embedded in online reviews may have a crucial impact on their perceived helpfulness. Reviewers often express their feelings freely in textual comments; in the case of seller reviews, this is particularly true when customers are especially pleased or displeased with their shopping experience. Given that review readers attend to both non-emotional and emotional aspects of a review, they are likely to perceive that embedded emotions are useful for understanding review content and making better decisions (Cao et al. 2011; Kuan et al. 2011). Following the conventional wisdom of a negativity bias, reviews with more negative emotions should be expected to be considered more helpful. However, this generalization from non-emotional to emotional aspects of seller reviews might not be adequate. Emotions are highly varied and complex, and cannot be reduced to simple ‘positive emotion’ and ‘negative emotion’ (Lerner and Keltner 2000). In particular, numerous different types of negative emotions occur in online reviews (anger, anxiety, disgust, etc.). Because these emotions are interpreted differently by readers, the effects of even same-valenced emotions may differ in systematic ways (Fontaine et al. 2007).

In order to address these issues, we ask the following research questions: How do reviewer emotions influence the perceived helpfulness of seller reviews? Specifically, does the impact of distinct emotions (such as anxiety and anger) vary in systematic ways, and why? Drawing from research in judgment and emotion, we propose that the effects of specific negative emotions vary due to underlying perceptions of reviewers’ cognitive effort. In particular, anxiety-embedded reviews are hypothesized to be considered more helpful than anger-embedded reviews, because anxious reviewers are expected to make more written effort and put more thought into the review than angry reviewers. Both field study with archival data and experimental methods are utilized to test these hypotheses.

We believe this paper to have important theoretical and practical implications. Emotions are extremely prevalent in online word-of-mouth, and we believe it is meaningful and important to explore their effects; ours is among the first attempts to examine the role of discrete emotions in the context of online reviews. Most importantly, by revealing that negative emotions differ from one another in consistent ways, we show that generalized, valence-based approaches may not be sufficient to fully explain the role of emotional content. Second, we explain the differential effect of two discrete negative emotions, anger and anxiety, on perceived review helpfulness. In doing so, we reveal that above and beyond the information content of a review, embedded emotions can impact perceptions of review helpfulness through inferences regarding the reviewer’s effort. Importantly, although our hypotheses only concern two specific emotions, the underlying arguments apply to discrete emotions in general (e.g., sadness, shame, disgust). Third, this paper integrates the advantages of both empirical and experimental methods. Given that both methods provide converging evidence for our hypotheses, this methodological triangulation enables us to be more confident in our findings. More practically, our findings stand to benefit reviewers, online retailers, and third-party review sites, by offering guidelines to help determine which reviews will be most influential. For example, our results suggest that reviews written by angry customers may not be as harmful as one might expect, and that retailers pay particular attention to anxiety-embedded reviews when managing their customer communication efforts. We close with a discussion of these and other implications.
Literature Review and Hypotheses

Information Diagnosticity and Emotions

As discussed in the preceding section, the helpfulness of a review is a reflection of the diagnosticity of the information it contains. Information diagnosticity (Feldman and Lynch 1988) is defined as “the degree to which one piece of information implies or determines one’s response to a given question or other circumstance requiring a judgment or behavior” (Feldman 1999, p. 48). In other words, a piece of information is diagnostic if it is informative for judgment and decision-making. Applied here, the diagnosticity concept aligns with prior research suggesting that seller reviews are perceived as helpful if they are diagnostic for evaluating a seller (Kuan et al. 2011; Mudambi and Schuff 2010).

In particular, abundant evidence supports the existence of a generalized negativity bias whereby “bad things will produce larger, more consistent, more multifaceted or more lasting effects than good things” (Baumeister et al. 2001, p. 325). One reason commonly cited for negativity bias is that negative information is less common than positive information and consequently perceived as more informative (Fiske 1980). Given that negative online feedback is much rarer than positive feedback, it is not surprising that negativity bias has been demonstrated within online word-of-mouth in e-commerce (see Pavlou and Dimoka 2006; Resnick and Zeckhauser 2002).

We argue that in addition to any negativity bias, the specific affective content embedded in online reviews also plays a major role in determining their diagnosticity. The term ‘affect’ is typically used to describe a general category of mental processing that reflects subjective internal feelings (Cohen et al. 2008). Both ‘moods’ and ‘emotions’ fall into the affect category, so it is helpful to distinguish between these two terms. Mood refers to a nonspecific, valenced feeling state that is typically low in arousal. In contrast, emotion refers to “a mental state of readiness that arises from cognitive appraisals of events or thoughts” (Bagozzi et al. 1999, p. 184). An emotion differs from a ‘mood’ in that an emotion tends to be briefer, more intense, more context specific, and more intentional with a particular cause (e.g., a dishonest seller) (Ekman 1992). Emotions have a specific, known source, and may lead to specific coping actions. Furthermore, many emotions are directly coupled with specific resulting action tendencies and behaviors. Although both mood and emotions undoubtedly play a role in the transmission of word-of-mouth, we focus on the emotion construct, because feelings expressed in seller reviews are targeted at particular purchase experiences and retailers.

Discrete Emotions

Work on the conceptual organization of emotions has consistently found valence to be an important dimension (Lang 1995). However, a fundamental weakness of the valence-based approach is that dimensions such as valence and activation/arousal are not useful for capturing emotions that do not differ in those dimensions (Fontaine et al. 2007). Given the variation and complexity of emotions (especially negative emotions), other differences may have nontrivial influence on the way that they are experienced and resolved. Therefore, this dimensional view of emotions has increasingly been challenged, and the result has been a much more nuanced exploration of distinct emotional states (Lerner and Keltner 2000).

In other words, even emotions with similar valence might have distinct impact on review helpfulness. The demonstration of this phenomenon requires evidence that specific emotions differ from each other in consistent ways (Levenson 1992), and this demonstration constitutes the primary contribution of our research.

Going beyond the valence-based approach, a variety of recent frameworks have been proposed for distinguishing the effects of discrete emotions on information processing and judgments (Lerner et al. 2007; Levenson 1992; Smith and Ellsworth 1985). One prominent approach theorizes that various emotions can be differentiated based on underlying dimensions of cognitive appraisal (Han et al. 2007). Under this approach, particular emotions result from specific patterns of appraisal regarding one’s environment (Smith and Ellsworth 1985). For instance, anxiety is characterized by appraisals of low certainty and situational control, whereas anger is characterized by high certainty and individual control. According to the appraisal-tendency framework (ATF) (Lerner and Keltner 2000; Lerner and Keltner 2001), emotions give rise to cognitive predispositions to appraise events in line with the appraisal
patterns that characterize the emotions, a process called “appraisal tendency” (Han et al. 2007). Thus, emotions also shape subsequent judgments (e.g., Berkowitz and Harmon-Jones 2004; Keltner et al. 1993), such as influencing judgments of review writers in terms of how much effort is required.

**Discrete Emotions and Cognitive Effort**

Dual-process theories of cognitive functioning distinguish between two modes of thinking and information processing: systematic (or central) processing and heuristic (or peripheral) processing (Bond et al. 2008; Chaiken and Trope 1999). The former refers to a “mindful” process that involves a great deal of thought and cognitive effort, whereas the latter refers to a “mindless” process that requires little thought or cognitive effort and relies on automatic cues or “rules of thumb”. Therefore, relative to heuristic processing, systematic processing is associated with heightened level of cognitive effort.

Integrating the theories above, we argue that emotions have predictable effects on the level of cognitive effort expended by the reviewer in writing their review, and the key determinant of these effects is the underlying appraisal of ‘certainty’ (Lerner and Tiedens 2006). Emotions associated with uncertainty appraisals (fear, anxiety, etc.) have been shown to induce systematic processing, whereas emotions associated with certainty appraisals (anger, disgust, etc) have been shown to induce heuristic processing (Tiedens and Linton 2001). As a result, uncertainty-appraised emotions tend to result in more cognitive effort than certainty-appraised emotions.

In the emotions literature, evidence for the ATF approach typically involves comparing emotions that are highly differentiated in certain appraisals on judgments that relate to those appraisal dimensions (Han et al. 2007; Lerner and Tiedens 2006). To the extent that judgment depends on the appraisals associated with an emotion rather than its valence, two ‘positive’ emotions may affect decisions quite differently, while a ‘positive’ and ‘negative’ emotion may have similar effects. In keeping with this approach, to illustrate the distinct effect of various emotions on cognitive effort and review helpfulness, we need to select emotions that vary in the certainty appraisal dimension. A sample illustration of emotions that differ in valence and certainty appraisal is shown in Table 1.

**Table 1. Sample Emotions Characterized by Valence x Uncertainty**

<table>
<thead>
<tr>
<th>Certainty</th>
<th>Valence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Surprise</td>
</tr>
<tr>
<td>High</td>
<td>Happiness</td>
</tr>
</tbody>
</table>

We focus on two specific emotions commonly encountered in seller reviews - anxiety and anger - that differ heavily in the certainty appraisal dimension. Following prior research, we conceptualize anxiety to subsume fear, frustration, worry (Brooks and Schweitzer 2011; Gray 1991), and conceptualize anger to subsume hate, disgust, irritation (Berkowitz and Harmon-Jones 2004; Shaver et al. 1987). Both anxiety and anger are negative emotions characterized by heightened activation/arousal (Barrett and Russell 1998), but anxiety is characterized by low certainty appraisal whereas anger is characterized by high certainty appraisal (Smith and Ellsworth 1985). In the domain of e-commerce, numerous antecedents can give rise to customer anxiety or anger (see later for specific examples). Anxiety in a seller review often stems from ambiguity regarding product quality, shipment times, or refunds/returns; anger often results from mishandled transactions, inadequate customer service, or poor product performance.

Due to their distinction in certainty appraisals, anxiety and anger should differ in their effects on the cognitive effort of the reviewer. Linked to uncertainty appraisals, anxiety activates systematic processing and results in increased sensitivity to threat and heightened vigilance to situational stimuli (Macleod and Mathews 1988; Yiend and Mathews 2001). In other words, anxiety is often associated with increased thoughtfulness and more careful processing of the threat inherent within a given situation (Huddy et al. 2007). In contrast, linked to certainty appraisal, anger activates heuristic processing and prompts a person to rely on superficial cues and more stereotypic thoughts (Bodenhausen et al. 1994; Tiedens 2001).
Thus, compared with anxiety, anger leads to less cognitive effort and less thorough processing of situational information (Lerner et al. 1998; Tiedens 2001).

In order to position our arguments, we differentiate perceivers (i.e., review readers) from actors (i.e., review writers). The arguments above imply that as actors, anxious reviewers will tend to engage in more systematic processing than angry reviewers. One determinant of review readers’ perceptions of review helpfulness is their assessment of writers’ cognitive effort, and we argue that this assessment is driven in part by the emotions embedded in the reviews. To make this argument, we begin by assuming that review readers are able to accurately identify discrete emotions in the content of seller reviews. Abundant evidence shows that people are able to judge various emotions fairly accurately from facial and bodily expressions (Atkinson et al. 2004; Ekman and Friesen 1971). In the realm of language, people can not only recognize emotional words (Zeelenberg et al. 2006), but also distinguish between various discrete emotions embedded in writing (Lindquist et al. 2006). Therefore, to the extent that a review contains emotional content, readers should generally be able to identify it. Next, we rely on substantial evidence that people can accurately associate emotion categories with corresponding cognitive appraisals (Scherer and Grandjean 2008). Therefore, readers’ naïve theories of emotional appraisal tend to (correctly) associate anger with greater certainty than anxiety, and as a result, readers will perceive greater cognitive effort from anxious reviewers than angry reviewers (Schwarz and Clore 2007). Finally, it is reasonable to expect that perceived cognitive effort of a reviewer will positively influence the perceived helpfulness of a review: based on the “more is better” heuristics (Chaiken 1987), the more effort reviewers are perceived to have expended in sharing their experience, the more diagnostic the reviews are likely to be considered.

Taking these arguments together, reviews written by anxious (vs. angry) consumers will result in a higher level of perceived cognitive effort, which will in turn lead to the perception that their reviews are more helpful. The theoretical framework is illustrated in Figure 1. Thus, we propose the following hypotheses:

Hypothesis 1: Anxiety-embedded reviews are perceived to be more helpful than anger-embedded reviews.

Hypothesis 2: Perceived cognitive effort mediates the differential impact of anxiety and anger on the perceived helpfulness of reviews.

In order to test these hypotheses, we conducted two studies utilizing distinct methods. In Study 1, we collected real seller reviews from Yahoo Shopping’s website and tested the impact of discrete emotions on review diagnosticity, operationalized by helpfulness ratings. Study 2 extended Study 1’s results using a controlled experiment. We manipulated anxiety and anger directly and explored the underlying process of their differential effects on perceived review helpfulness.

**Study 1: Yahoo Reviews**

The primary goal of Study 1 was to explore the effects of discrete emotions on review helpfulness. To do so, we collected and analyzed actual review data from the Yahoo Shopping website, which provides both user ratings and text reviews for online merchants. The review page of each merchant displays all its reviews chronologically, and the most recent reviews appear first by default. Prior consumers of a retailer can evaluate that retailer by leaving a rating on a scale of 1 to 5 stars. Additionally, they can write a text
review to provide more details about their experience with the merchant (Yahoo!), as illustrated in Figure 2.

![Figure 2. Study 1: Screenshot of a Yahoo Retailer Review](image)

**Data Collection**

Data collection took place in April 2011. We began by targeting stores in the “Electronics” category, which sell a wide range of products (e.g., cameras, cell phones, televisions, MP3 players, home video, etc.). We retrieved all historical reviews for over 8 years; merchants without a single review were eliminated, leaving 167 stores. For each of these stores, we collected the following information regarding each review: rating, text review content, helpful votes, and total votes. We also collected store-level information, such as the average rating and count of all ratings of each store. Individual reviews were the unit of analysis.

In order to reduce noise in the reviews, the following steps were taken. First, 562 reviews included non-ASCII characters (mostly from non-English languages), and these reviews were dropped. Next, we dropped reviews that contained no text content (4,571), reviews that contained only “EOM” (representing “End of Message”, 27,708), and reviews that contained only symbols or dates (10). These steps resulted in 154,834 reviews. Of this set, only 7,322 reviews (4.7%) were ever voted on, which is not unusual in online review settings (Pavlou and Dimoka 2006). Analysis was conducted on these 7,322 reviews.

**Variables**

The dependent variable of interest, review helpfulness, was operationalized as follows. Below each review, Yahoo Shopping asks the question “Was this review helpful?” with “Yes” and “No” options. A review that has received at least one vote will display the number of “helpful” votes and total votes immediately before the review content. Helpfulness was measured as the proportion of “helpful” votes out of the total votes a review received (i.e., the number of people who voted “Yes” divided by the total number of people who cast a vote). A review is on average relatively helpful if the percentage of “helpful” votes is high. The value of helpfulness for voted reviews ranged from 0 to 1. The average helpfulness of the voted reviews was 0.68, indicating that most reviews were considered relatively helpful. Tables 2 and 3 present a summary of statistics and correlations for this and the remaining variables (described below).

| Table 2. Study 1: Descriptive Statistics for Voted Reviews (N = 7,322) |
|---|---|---|---|---|
| Variable | Mean | Std. Dev. | Min | Max |
| 1 Review helpfulness | 0.68 | 0.40 | 0 | 1 |
| 2 Rating | 3.29 | 1.81 | 1 | 5 |
| 3 Length | 69.82 | 70.76 | 1 | 707 |
| 4 Reading difficulty | 10.32 | 4.25 | -10.2 | 121.5 |
| 5 Anxiety | 0.17 | 1.00 | 0 | 50 |
| 6 Anger | 0.19 | 1.14 | 0 | 50 |
The independent variables of interest include the presence of two discrete emotions: anxiety and anger. To measure these variables, we utilized Linguistic Inquiry and Word Count (LIWC), a text analysis software program designed by Pennebaker, Booth, and Francis (2007). LIWC maintains a dictionary composed of almost 4500 words and word stems, each of which defines one or more categories. Importantly, LIWC has categories that correspond to anxiety and anger. LIWC processed each individual text review, one word at a time, and computed scores for anxiety and anger using a matching procedure. As each word in the review was processed, LIWC searched its dictionary file for a match with the review word. If a match occurred, the appropriate category scale(s) for that word would be incremented. At the end of this procedure, a final score was assigned for each category, representing the percentage of words in the review that matched that category. Although the value of anxiety and anger scores could range from 0 to 100, the maximum value for voted reviews in our data set was 50, and the average scores for both anxiety and anger were below .2. In fact, 9.81% of voted reviews contain at least one anxiety word identified by LIWC, and 11.84% of voted reviews contain at least one anger word. This is not surprising given that LIWC can only identify emotional words that match its pre-defined dictionary. Some examples of anxiety-embedded and anger-embedded reviews are listed in Table 4.

<table>
<thead>
<tr>
<th>#</th>
<th>Anxiety-Embedded Reviews</th>
<th>Anger-Embedded Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I had some doubts about the item I purchased, never got an answer neither the store nor the manufacturer</td>
<td>Lied about availability of product for two weeks, indicating that it had been shipped when, in fact, it was on back-order. Customer service? Don’t bother!</td>
</tr>
<tr>
<td>2</td>
<td>Lost order per customer representative. No explanation. Now I am worried that they will &quot;find&quot; the order and will have to return since I am ordering from another vendor.</td>
<td>These people SUCK. They stalled the order for days trying to get me to buy extra shipping and other crap. Then they screwed up and didn’t ship me one of the TV’s I ordered. They SUCK.</td>
</tr>
<tr>
<td>3</td>
<td>The product was &quot;backordered&quot;. It was ordered over a month ago as a gift, good price but never received the item. Said they would refund my credit card in 72 hours, and it’s been over a week and no refund. Getting a little worried. They are quick to reply to e-mails, but no refund. Seems to be a good company on yahoo, will update if the refund is made. (4th of July Holiday)</td>
<td>Extremely disappointed and offended. My Miele machine broke after 10 uses. When I called the store today, I was told that I was an idiot and that I was wasting 11 minutes of the salesperson’s time with my idiocy. Then he hung up on me. I am contacting Miele headquarters to complain as well. I will never do business with this store again, and if you don’t want to get ripped off and abused, you shouldn’t either.</td>
</tr>
</tbody>
</table>

Our analysis controlled for a series of relevant variables, including rating and rating squared, review length, review reading difficulty, and certain store characteristics. (1) Rating refers to the star rating of a review; the more stars a review receives, the more positive the review is. Rating ranged from 1 star to 5 stars, and the average rating for the voted reviews was 3.29. (2) A quadratic term of star rating was included to account for the non-linear relationship between rating and helpfulness (Mudambi and Schuff...
Review length was operationalized as the number of words in a review; a longer review often provides more useful information, thus considered more helpful. The voted reviews had on average 69.82 words. (4) To control for the understandability of reviews, we calculated the Coleman–Liau Index, an estimate of the U.S. grade level that a student would need to have achieved in order to read and understand the text (Coleman and Liau 1975). On average the voted reviews were written at a 10th grade level. Lastly, we controlled for the effects of store characteristics, including a store’s average rating and the count of all its prior ratings. The former captures the overall reputation of a store, while the latter captures popularity. The operationalization of all variables is summarized in Table 5.

Table 5. Study 1: Variables and Operationalizations

<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable Level</th>
<th>#</th>
<th>Variable</th>
<th>Operationalization</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV</td>
<td>Individual Review</td>
<td>1</td>
<td>Review Helpfulness</td>
<td># helpful_votes / # total_votes</td>
<td>Range: [0, 1]</td>
</tr>
<tr>
<td>IV</td>
<td>Individual Review</td>
<td>2</td>
<td>Anxiety</td>
<td>(# anxiety-related words / # words in a review) * 100</td>
<td>Range: [0, 100]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Anger</td>
<td>(# anger-related words / # words in a review) * 100</td>
<td>Coded by LIWC</td>
</tr>
<tr>
<td>Control</td>
<td>Individual Review</td>
<td>4</td>
<td>Rating</td>
<td># of stars</td>
<td>Range: [1, 5]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Length</td>
<td># of words</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Reading Difficulty</td>
<td>Coleman-Liau Index</td>
<td>U.S. grade level necessary to comprehend the text</td>
</tr>
<tr>
<td>Store</td>
<td></td>
<td>7</td>
<td>Reputation</td>
<td>average rating</td>
<td>Range: [1, 5]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>Popularity</td>
<td># of ratings in total</td>
<td></td>
</tr>
</tbody>
</table>

Data Analysis and Results

Because review helpfulness is a limited dependent variable, we followed the approach of Mudambi and Schuff (2010) by using Tobit regression for data analysis. We deemed this analysis appropriate for a number of reasons. First, the dependent variable was censored in nature: its value was bounded in range because it was constructed as a ratio. Additionally, prior literature has shown that the proportion of reviews receiving helpfulness votes varies with website popularity and review targets (Cao et al. 2011), and in our data set the dependent variable was actually missing for most of the reviews. Second, there exists a potential selection bias, because not every review reader casts a helpfulness vote. More importantly, the probability of a review being voted on might be correlated with explanatory variables such as review rating, review length, etc. Therefore, a sample containing only voted reviews might be non-random, and least-squares estimation of this sample would produce biased estimates (Greene and Zhang 2003). Thus, we utilized a Tobit model to analyze reviews with at least one vote (N = 7,322).

Table 6 contains the results of our empirical analysis. In the analysis, we standardized all independent variables to unify presentation of the results. The analysis indicates a good fit, with a highly significant likelihood ratio ($p < 0.001$) and a pseudo $R^2$ value of 0.239 (Veall and Zimmermann 1996).
Tobit regression results concerning the control variables were largely consistent with prior literature. The linear (β = -2.196, \( p < 0.001 \)) and squared (β = 1.859, \( p < 0.001 \)) coefficients of review rating were significant and in the expected direction: reviews with lower ratings and higher extremity were considered as more helpful. Additionally, a review was considered as more helpful to the extent that it was longer (β = 0.207, \( p < 0.001 \)), and less difficult to understand (β = -0.152, \( p < .001 \)). The coefficients for average rating (β = -0.605, \( p < 0.001 \)) and count of ratings (β = -1.080, \( p < 0.001 \)) were significant and in the expected direction; that is, controlling for all other variables, if prospective customers encounter a reputable and popular store, its reviews are considered less diagnostic.

To test our first hypothesis, we compared the coefficients of the two discrete emotions. The coefficient of anxiety on review helpfulness was positive and significant (β = 0.034, \( p < 0.05 \)). In contrast, the coefficient of anger was negative but not significant (β = -0.013, \( p = 0.360 \)). Most importantly, a comparison of the two discrete emotions’ coefficients revealed that their difference was significant (F(1, 154826) = 5.86, \( p < 0.05 \)), supporting H1. As expected, adding an anxiety-related word to a review increased helpfulness to a greater degree than adding an anger-related word.

Discussion

Utilizing actual review data from Yahoo Shopping websites, this study provided initial evidence for our first hypothesis. In the context of seller reviews, words related to anxiety differed from those related to anger in terms of their impact on review helpfulness.

However, the empirical methods that were utilized necessitated several limitations. Although our Hypothesis 2 proposes that emotion-based differences in perceived helpfulness are driven by the perceptions of cognitive effort, no reasonable measure of this mediator was available. Therefore, other explanations cannot be ruled out. Most prominently, the results might simply result from the fact that anxious reviewers actually do produce more helpful content than angry reviewers. Although this possibility and our theory are not mutually exclusive, a stronger test of H1 necessitates an approach which controls for informative review content. Moreover, the empirical method used in the study could not provide definitive evidence for causality because the method was cross-sectional. To address these limitations, we conducted a controlled experiment in Study 2.
Study 2: Experiment

In this study, we utilized a laboratory experiment to directly manipulate anxiety and anger in seller reviews, while holding constant any substantive content. Seventy-eight undergraduate students from a southern U.S. university participated in the study. In a simulated feedback scenario, each participant was exposed to reviews of three potential stores: one written by an anxious reviewer, one by an angry reviewer, and one by a non-emotional reviewer. By comparing the perceived helpfulness across stores, we could identify the differential impact of anxiety and anger on review helpfulness. In addition, we measured the process variable of perceived cognitive effort and tested its mediating effect.

Stimulus Materials

Using the real reviews collected in Study 1 as a starting point, we developed a set of 13 non-emotional text reviews. First, we generated a pool of one-star negative reviews, and revised the reviews by removing any sentences which directly indicated reviewer emotions. Of these, we selected 13 reviews containing content that could reasonably have been written by both anxious and angry customers.

Next, we conducted pretests on this set to identify reviews that were perceived to reflect equivalent levels of anxiety and anger. 25 pretest subjects were asked to read the 13 reviews in turn, and rate the anxiety and anger of each review writer. The three reviews selected for the experiment are shown in Table 7. For each review, the difference between anxiety and anger evaluations was not significant ($p > 0.8$).

Table 7. Study 2: Review Stimuli Used in Experiment

<table>
<thead>
<tr>
<th>#</th>
<th>Text Review Content</th>
<th>Anxiety</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I purchased a camera on Feb 27th for two day delivery and on March 23 I am still waiting for it, plus they billed me for it on Feb 27th.</td>
<td>6.7</td>
<td>6.8</td>
</tr>
<tr>
<td>2</td>
<td>Ordered a laptop battery (12 cell) and RAM. I received a 6 cell battery and the incorrect RAM. I returned the products to this merchant three weeks ago (and they were received), but still have not received my refund.</td>
<td>6.8</td>
<td>6.8</td>
</tr>
<tr>
<td>3</td>
<td>I placed an order on Dec. 14th using standard shipping because it said if I ordered by the 19th it would be delivered before Christmas. I just received an E-mail saying they shipped it today (Dec. 23rd) and estimated arrival date is Dec. 30th.</td>
<td>6.5</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Procedure

The cover story introduced a fictitious third-party review site, “OnlineConsumerReview.com,” that provides consumer reviews of online stores. Participants were told that the experimenters had been working with this site to improve its data mining algorithms. As a result, they would be asked to evaluate a series of real text reviews from the site.

Participants read and evaluated 6 text reviews, one at a time, each describing a different online store, 3 positive ‘filler’ reviews were presented in positions 1, 3, and 5, and the 3 treatment reviews were presented in positions 2, 4, and 6. The sequence of the three conditions was counterbalanced so that each review was paired with each condition an equivalent number of times.

The treatment reviews are presented in Table 7. Emotion was manipulated directly by varying the first sentence of the review. In the anxiety condition, the review began with the sentence “My experience with this seller has caused a lot of anxiety.” In the anger condition, the review began with the sentence “I was very angry after everything that happened”. The review in the baseline (control) condition contained no additional sentence. In all other respects, reviews across the three conditions were identical.

After reading each review, participants reported their perceptions of: 1) the helpfulness of the review, and 2) the cognitive effort expended by the reviewer. Perceived review helpfulness was measured on a 9-point semantic differential scale, using three items adapted from Sen and Lerman (2007). Perceived cognitive
effort was measured on a 9-point scale ranging from “not at all” to “very much”, using three items adapted from Huddy, Feldman and Cassese (2007). These measures are presented in Appendix.

**Results**

Before further analysis, we conducted a manipulation check of the stimulus materials. A separate set of 30 subjects went through a procedure similar to the main study, but the dependent measures after each review were replaced with the following question regarding emotions: “In your opinion, to what extent does each of the following words describe how the reviewer felt when he/she wrote the above review?” The options include “anxious” and “angry,” and each option contained a 9-point scale (1 = “Not at all” and 9 = “Very much”). Analysis was performed using pairwise comparisons after a repeated-measure ANCOVA (controlling for the order of reviews). Confirming that the treatment reviews successfully targeted their relevant emotions, reviews in the anxiety condition were more related to anxiety than to anger ($M_{anxiety} = 8.267$ vs. $7.200$, $p = 0.013$), and reviews in the anger condition was more related to anger than to anxiety ($M_{anger} = 8.700$ vs. $6.267$, $p < 0.001$). Additionally, reviews in the control condition were related to both anxiety and anger to a similar extent ($M_{control} = 6.867$ vs. $7.167$, $p = 0.344$).

We also examined the reliability and validity of major constructs in the study. For each of the three treatment reviews, Cronbach’s alphas for both constructs were well above 0.80, demonstrating adequate internal consistency reliability (Nunnally 1967). Next, we conducted an exploratory factor analysis (EFA) to assess convergent and discriminant validity of the two constructs. The principle components method with Varimax rotation was used. For each review, EFA consistently provided two factors. Moreover, in the rotated component matrix: loadings of items on their corresponding factor were higher than 0.7, higher than loadings of other items on this factor, and higher than the loadings of these items on the other factor ($< 0.5$) (Straub 1989). Together, these results indicated adequate convergent and discriminant validity.

The first important question concerns whether perceived helpfulness varied across anxious vs. angry reviews. The pattern of means is illustrated in Figure 3. A repeated-measure ANCOVA was performed to examine the difference in perceived helpfulness across treatment reviews. The emotional condition was entered as a within-subject factor, and the counterbalancing of the three treatment reviews was entered as a covariate. In line with $H_1$, pairwise comparisons revealed that the difference in perceived helpfulness between anxiety and anger conditions was significant ($M_{anxiety} = 7.57$ vs. $7.23$, $p < 0.05$). Thus, reviews written by anxious customers were considered more helpful than those written by angry customers.

![Figure 3. Study 2: Perceived Helpfulness of Seller Reviews Across Emotion Conditions](image)

As a supplementary analysis, we also compared the helpfulness of emotional reviews with that of the baseline review. Pairwise comparisons showed that anxious reviews were considered significantly more
helpful than baseline ($M = 7.57$ vs. $7.00$, $p < 0.001$), whereas angry reviews were not ($M = 7.23$ vs. $7.00$, $p = 0.16$). Taken together, these results indicate that negative reviews were considered more helpful if they indicated anxiety but not if they indicated anger.

Next, we explored whether the differential effects of anxiety and anger on perceived helpfulness were mediated by perceived cognitive effort. The pattern of means is illustrated in Figure 4. When perceived cognitive effort of reviewers was entered as the dependent variable, a repeated-measure ANCOVA showed results similar to those above. Specifically, the difference in review helpfulness between anxiety and anger conditions was significant ($M = 6.27$ vs. $5.82$, $p < 0.01$); the difference between anxiety and control conditions was significant ($M = 6.27$ vs. $5.46$, $p < 0.001$), and the difference between anger and control conditions was significant ($p < 0.05$). In sum, anxious reviewers were perceived to have spent more effort than angry reviewers in writing the review.

We next examined the two criteria recommended by Judd, Kenny and McClelland (2001) for testing mediation in within-subject designs. First, the difference in perceived cognitive effort between anxiety and anger reviews was significant ($M = 6.27$ vs. $5.82$, $p < 0.01$) and in the same direction as the difference in perceived helpfulness. Second, the difference in perceived cognitive effort between anxiety and anger reviews was predictive of the difference in perceived helpfulness ($p < 0.001$). Therefore, perceived cognitive effort mediated the differential impact of anxiety and anger on the perceptions of review helpfulness.

**Discussion**

By directly manipulating discrete emotions and measuring perceived cognitive effort, Study 2 provided a replication and extension of Study 1. The main advantage of the experimental method was the straightforward manipulation of anxiety and anger while avoiding extraneous factors. This parsimony enabled us to explore the reasons for differential effects of anxious and angry reviews. Results converged with those of Study 1 and with our theoretical arguments: participants considered anxious reviews to be more helpful than angry reviews (although information in the reviews was actually identical), and this difference was mediated by the perceived cognitive effort of the reviewer. Both H1 and H2 were supported.
General Discussion

Together, the empirical investigation in Study 1 and experimental exploration in Study 2 provided converging evidence for our hypotheses. Moving beyond valence (‘good’ and ‘bad’) to focus on discrete emotions, we demonstrated the differential impact of anxiety and anger on review helpfulness: reviews written by anxious reviewers were considered more helpful than those written by angry reviewers. Furthermore, this differential impact on perceived review helpfulness can be explained by perceptions regarding the cognitive effort of reviewers.

Theoretical Implications

This paper provides multiple contributions to existing literature in online word-of-mouth and emotions. First, in contrast to the current, cognition-dominated literature on review helpfulness, this paper is among the first attempts to explore the effects of emotions above and beyond their non-emotional counterparts. Within the field of information systems more generally, emotions and affective processes are frequently overlooked. To explain the effectiveness of information and communication technologies, almost all prominent IS theories and conceptual frameworks left out the emotional components, such as Technology Acceptance Model (Davis 1989), Media Richness Theory (Daft et al. 1987), and Task Technology Fit theory (Goodhue and Thompson 1995). On the other hand, management scholars in recent decades have begun to recognize the importance and ubiquitous role of emotions in people's decision-making processes (Loewenstein and Lerner 2003), and research dealing with emotions has exploded especially in marketing (Bagozzi et al. 1999; Cohen et al. 2008) and organizational behavior (Ashkanasy et al. 2002; Brief and Weiss 2002). While IS studies examining emotions are still relatively scarce, some researchers have started to investigate the impact of emotions in various contexts. For example, IT adoption scholars incorporated perceived affective quality (Zhang and Li 2005), perceived enjoyment (Sun and Zhang 2006; Yi and Hwang 2003), and computer anxiety (Venkatesh 2000) into the user technology acceptance framework. Scholars examining web interface designs observed the effects of a web user's initial emotional responses on subsequent behaviors (Deng and Poole 2010). Scholars on trust in e-commerce demonstrated the mediating effects of website users' emotions on the development of online trust (Hwang and Kim 2007). Recognizing the importance of studying emotions in IS research (see de Guinea and Markus 2009), we believe that our research conducted in the online word-of-mouth context deepens the understanding of emotions in the online environment and contributes to this burgeoning area.

Second, we disentangled the differential roles of two discrete emotions in evaluations of review helpfulness. Extending the logic of negativity bias to the emotion context, one would assume that reviews with negative emotions would be more helpful. However, this valence-based approach cannot explain the distinct effect of various emotions similar in valence (Fontaine et al. 2007). To demonstrate this point, we concentrated on two specific emotions: anxiety and anger. Both of these are negative, highly arousal emotions; nevertheless, due to their distinct motivations and behavioral implications, we expected and observed that they would influence perceptions of review helpfulness in distinct ways. There have been calls to move ‘beyond valence’ in examining the effect of emotions (Lerner and Keltner 2000), and some IS scholars have explored the distinct roles of discrete emotions in technology acceptance (Venkatesh 2000) and online trust (Hwang and Kim 2007). In line with this movement, our studies provided initial evidence that emotional reviews - even of the same valence - are not all created equal regarding perceptions of review helpfulness.

Third, we examined the mechanism underlying the two emotions’ differential impact. Existing IS literature in this area often treated emotions as a mediator, assuming that positive emotions would result in positive outcomes, and vice versa for negative emotions. In contrast, we probed beyond valence and provided more nuanced explanations with regard to why emotions matter in determining the diagnosticity of review information. Specifically, building on dual-process theories and the categorization of discrete emotions, we proposed perceived cognitive effort as a mediator, and showed its mediating role through an experimental study. In so doing, this paper provided useful evidence for the appraisal-tendency framework, which claims that emotions differing in certainty appraisals are associated with distinct thought processes (Lerner and Tiedens 2006; Tiedens and Linton 2001). Although we focused on two specific emotions in this paper, the same logic could be applied to other emotions to predict their
distinct effects in online word-of-mouth. More importantly, the underlying mechanism could be used to explain the effect of discrete emotions on other important dependent variables such as perceptions of reviewer expertise and trustworthiness judgments of retailers.

Fourth, we explored the perceptions of emotions rather than emotions per se. Relevant research in psychology and related disciplines tends to examine the impact of discrete emotions in settings involving only one or two parties (Van Kleef et al. 2010). This paper expanded the setting to three parties, exploring how prospective consumers make sense of emotions expressed by prior buyers toward online retailers. In particular, review writers’ emotions may not influence review readers directly. Instead, the communicated feelings provide useful information, above and beyond that of ratings and objective review content, by which prospective customers may make spontaneous inferences about the reviewers (Dick et al. 1990; Naylor et al. 2011). In this paper, we observed that naive theories of review readers about the construction of reviews impacted perceptions of their value. Specifically, perception of reviewers’ cognitive effort, which was inferred from the emotions conveyed in the review content, mediated the effect of discrete emotions on perceived review helpfulness. We believe that this paper advanced our understanding of emotions occurring in the real world by introducing the perspective of actors versus perceivers and focusing on the observers’ perceptions.

Fifth, online word-of-mouth in e-commerce is much more complicated than mere ratings. Going beyond ratings and reviewer characteristics that were easily observable in prior empirical investigations, we utilized content analysis and experimental method to demonstrate that emotions inferred from the textual content of seller reviews can predict the helpfulness perceptions of the reviews. Consistent with the implications of other papers looking into the substance of reviews themselves (Cao et al. 2011; Pavlou and Dimoka 2006), our findings suggested that the rich contextual information from text reviews and the embedded emotions are useful in explaining what constitutes a helpful review.

### Practical Implications

Our findings have implications for review writers, who may have incentives (e.g., altruism, self recognition) to write more “helpful” reviews to inform future consumers. When writing a text review, an unsatisfied customer often express negative emotions, which have the potential to influence the attitude and behaviors of future customers to a greater extent than positive emotions. However, this does not mean that the more emotional a negative review is, the more helpful. Instead, this conclusion should be qualified according to the types of emotions involved. For instance, our findings about the differential effects of anxiety and anger suggested that an anger-embedded review is less helpful than an anxiety-embedded review even if the substantial content of the review is held constant; anger will be associated with less cognitive effort than anxiety in review readers’ minds. As a result, ranting about a bad experience may not be a good strategy for reviewers who care about the quality of their produced information. To generate more helpful reviews, reviewers could postpone the writing for some time so as to allow them to calm down, think more rationally, and put more actual effort in the writing. It may also be beneficial if reviewers could take the perspective of readers rather than simply expressing their feelings to satisfy their own needs. Alternatively, they could try to avoid or at least suppress explicit expressions of their anger in a review, which might compromise the perceived quality of a review even if the review contains helpful information.

Our findings may also benefit online retailers seeking to utilize seller reviews to boost trust and increase sales. Various empirical studies have explored the helpfulness of product reviews (Chevalier and Mayzlin 2006; Forman et al. 2008; Mudambi and Schuff 2010), providing implications for product manufacturers. This paper extended the investigation to seller reviews, which are of particular concern to online retailers. Consistent with suggestions based on a negativity bias, retailers need to be especially vigilant of negative reviews. From the perspective of actors versus perceivers, an unsatisfied customer himself/herself may mean nothing for a seller, but their written reviews and expressed emotions may have detrimental effect on hundreds of thousands of potential newcomers. More importantly, reviews are helpful in the sense that they facilitate the purchase decision process, so retailers need to take special care of those negative reviews that are more helpful and thus more impactful for prospective consumers. Common sense indicates that an angry reviewer would be more undesirable than an anxious reviewer. Surprisingly, our findings suggest that reviews written by angry reviewers may be less of a concern for sellers, because they will be considered less helpful by prospective readers. However, reviews containing
anxiety or other uncertainty-appraised emotions are more influential on the behaviors of future customers; they may need to be taken care of promptly and properly. For example, many third-party review sites (e.g., BizRate.com) provide the functionality for retailers to post a verbal response/comment publicly right below a review, and retailers can benefit by taking advantage of these opportunities. In the response, retailers can probably make sincere apologies, explain the situation clearly, and clarify any misunderstandings so as to reduce the potential harm caused by these emotional reviews.

Lastly, third-party review sites can also gain by improving the site design based on our findings. First, some review platforms do not allow sellers to post response for a negative review, such as Yahoo! Shopping. Implementing this functionality could not only benefit retailers who can clarify misunderstandings and remedy their mistakes, but also boost the overall information quality of the website accordingly. Second, review sites can develop writing guidelines to encourage more useful seller reviews and ensure greater site stickiness. Many websites do not have any guidelines for writing a helpful review, such as eBay and Yahoo! Shopping. Some other stores listed guidelines beside the writing window. For instance, Epinions.com has guidelines such as “Do not use offensive language or content”, which is consistent with our implications regarding anger. However, this site has no guidelines targeted specifically at emotions. Third-party review sites can probably ask reviewers to feel free to express emotions, but should encourage them to think more carefully in general before the writing and avoid offensive or ranting language.

**Future Research**

A number of unexplored questions warrant future research, and we focus here on two in particular. First, the studies in this paper were conducted in the seller-review context. Although there is good reason to assume that the same arguments and underlying processes will apply to product reviews, additional factors need to be considered. For instance, the party reviewers are emotional towards is less clear with products than it is with sellers, limiting the potential impact of emotions. However, reviewers of sellers are mostly anonymous, whereas reviewers of products can generally be identified in terms of their expertise, location, etc. Future research can explore whether these variables strengthen or weaken the effect of discrete emotions.

Second, we explored the differential effects of two discrete emotions: anxiety and anger. Although these two emotions are very prevalent in seller reviews, other emotions also occur, such as sadness, regret, etc. It is worth considering whether cognitive efforts also mediate the effect of these other emotions on review helpfulness, and what other mediating processes may be responsible. For instance, another commonly observed negative emotion in online word-of-mouth is disappointment, which is characterized by higher certainty (van Dijk and Zeelenberg 2002). Based on our arguments, reviewers who are disappointed will be perceived to have expended less effort than those who have low-certainty emotions; accordingly, disappointment-embedded reviews will be perceived less helpful.

**Appendix: Variables Measured in the Experiment (Study 2)**

**Helpfulness:**
Using the scales below, how would you describe the above consumer review?
- not at all helpful / very helpful
- not at all useful / very useful
- not at all informative / very informative

**Perceived cognitive effort of reviewers:**
- In your opinion, how much effort had the reviewer put into writing this review?
- In your opinion, how much thought had the reviewer given to the above review when he/she wrote it?
- In your opinion, how much time did the reviewer spent writing this review?
References


