
Completed Research Paper

Chih Hung Peng
City University of Hong Kong
Tat Chee Avenue 83, Kowloon, Hong Kong SAR
chpeng@cityu.edu.hk

Dezhi Yin
University of Missouri
Cornell Hall, Columbia, Missouri 65211
yind@missouri.edu

Chih-Ping Wei
National Taiwan University
Roosevelt Rd., Taipei City 106, Taiwan
cpwei@im.ntu.edu.tw

Han Zhang
Georgia Institute of Technology
Tech Square, Atlanta, GA 30332
han.zhang@scheller.gatech.edu

Abstract:

Although longer reviews are generally considered more helpful, no research has investigated whether “the more the better” also applies to the expression of emotions. This paper explores the distinct effects of review length and emotional intensity. We propose that, in contrast to review length, the intensity of emotions has a negative effect on review helpfulness, and that this effect only applies to positive emotions. Additionally, drawing on elaboration likelihood model and the literature on the social functions of emotions, we predict that the respective effects of review length and emotional intensity are moderated by reviewer trustworthiness and the difficulty of reading review content. To test these hypotheses, we collected a rich data set from Epinions.com - a leading provider of consumer reviews. Our findings reveal the importance of taking the intensity of emotions into consideration when evaluating review helpfulness, and the results carry important practical implications.

Keywords: E-business, online review, review helpfulness, emotion, emotional intensity, empirical analysis
Introduction

Since the inception of electronic commerce, online reviews have played an extremely important role in consumer decision-making. It has been a common practice for retailers (e.g., Amazon, Home Depot, Walmart, etc.) and product manufacturers (e.g., Hewlett-Packard, Dell, Nike, etc.) to create and maintain their own online review platforms for consumers to communicate their opinions about certain products (Ku et al. 2012). However, given the abundance of reviews available online, consumers cannot afford to go through thousands of reviews about the products of their interest. Therefore, many online vendors provide various mechanisms to identify online reviews that customers perceive as most helpful (Cao et al. 2011; Mudambi and Schuff 2010). The evaluation of review quality and identification of the most helpful ones have become increasingly critical (Yin et al. 2014).

In recent years, we have witnessed a proliferation of studies focusing on online reviews and their helpfulness. Online reviews refer to peer-generated evaluations posted on company or third party websites. Following prior research, we define review helpfulness as the extent to which a peer-generated evaluation is perceived by consumers to facilitate their purchase decision process (Mudambi and Schuff 2010). Prior literature examining antecedents of review helpfulness has focused on factors such as review length, numerical ratings, product category, timing, and reviewer characteristics (Chen and Lurie 2013; Forman et al. 2008; Mudambi and Schuff 2010; Pan and Zhang 2011). One of the central concerns in this paper is the length of a review. Because review length reflects the amount of information and the effort of reviewers, studying its impact and boundary conditions has critical implications for reviewers who strive to craft helpful reviews and for consumers who make their purchase decision based on reviews. Consistent with a common understanding of “the more the better,” review length has been found to positively influence review helpfulness (Mudambi and Schuff 2010; Yin et al. 2014). Mudambi and Schuff (2010) also demonstrate that product type moderates the effect of review length. However, whether the positive effect of review length is contingent on source and other content characteristics is still relatively unknown.

Moreover, a fundamental characteristic of online review, its emotional content, has rarely been explored. Among a limited number of studies exploring emotional content, Chen and Lurie (2013) and Kuan et al. (2011) demonstrate that perceived helpfulness of a review is impacted by its emotional valence. Yin et al. (2014) study the impact of two specific emotions (anxiety and anger) on perceived helpfulness. However, no research has examined how the intensity of emotions relative to non-emotional content in a review could shape consumer evaluations of review helpfulness.

In this research, we ask the following research questions: How do the length of a review and the intensity of emotional expressions affect review helpfulness, respectively? In addition, are these relationships moderated by source and content characteristics, and how? First, we hypothesize that review length has a positive effect on review helpfulness because a longer review contains more details in the review and reflects more reviewer effort. Second, we predict that the intensity of emotions, especially positive emotions, has a negative effect due to readers’ suspicion of review manipulation and online fraud. Third, drawing on elaboration likelihood model and the emerging literature on the social functions of emotions, we further predict that the respective effects of review length and emotional intensity are moderated by the trustworthiness of reviewers and the difficulty of comprehending the review content.

To test our hypotheses, we utilized a rich data set collected from Epinions.com. Epinions.com is a premier consumer reviews platform on the Internet, which covers millions of products and services in over 30 different categories. On its website, people can search for products, read reviews and ratings from other people, and contribute their own reviews to share their good or bad experiences. Our empirical analyses reveal that a review will be rated more helpful if the review is longer or if the review has a lower level of emotional intensity. Moreover, we find that both reviewer trustworthiness and review reading difficulty moderate the aforementioned relationships.

Given a lack of understanding of the impact of emotional content on consumer behavior, our paper has important implications for reviewers, product manufacturers, and review platforms. Although a longer review is in general more helpful, gauging the expression of emotions is a constant challenge for reviewers striving to write the most helpful reviews. On one hand, expressed emotions may provide useful information (Yin et al. 2014). On the other hand, readers may perceive emotions negatively as signals of review manipulation and online fraud (Xiao and Benbasat 2011). In addition, the respective effects of
Review length and emotional intensity can be mitigated if the reviewer is more trustworthy or if the review content is written in a sophisticated manner. This research deepens our understanding of how and when review content, especially the intensity of emotions, affects review helpfulness. The rest of the paper is organized as follows. Section 2 reviews the related literature and proposes our hypotheses. Section 3 presents the research methodology and discusses the findings. Section 4 concludes the paper with theoretical and practical implications.

Hypotheses and Theory Development

Elaboration Likelihood Model and Review Length

Elaboration likelihood model (ELM) provides the theoretical foundation for our first set of hypotheses related to review length. The impact of online reviews on review readers could be described as a persuasion process, whereby potential consumers form or modify their attitudes about a product based on the review content. According to ELM, a persuasion process can involve different amount of thinking along an “elaboration continuum” (Petty and Briñol 2012). At the high end of the continuum, persuasion is said to follow a central route whereby people engage in extensive thinking and elaboration of messages. In contrast, at the low end of the continuum, persuasion is said to follow a peripheral route whereby people put relatively little thought into the message but rely heavily on peripheral cues, such as source trustworthiness. Most of the time these two routes coincide and jointly influence persuasion (Petty et al. 1981). However, when people are motivated and able to process a message, the likelihood of elaboration (i.e., the extent to which people elaborate on the message content) tends to be high.

The length of a review is a key determinant of review helpfulness as consumers take the central route of reading reviews carefully (Baek et al. 2012). At online review sites, readers are in general interested in the reviews that may assist them in evaluating alternatives and making purchase decisions, so they tend to follow the central route. As they scrutinize the content of a review, readers are likely to appreciate the details contained in it and the effort its reviewer devotes to crafting the review (Latif 2009; Ma et al. 2013). All else being equal, longer reviews presumably contain more details such as the pros and cons of a product, justifications of personal opinions, and intricacies of using the product in specific contexts. The persuasion literature finds that a larger number of reasons and arguments are more persuasive and they lead to more confidence in the decisions (Luo 2002; Maddux and Rogers 1980; Schwenk 1986; Tversky and Kahneman 1974). Similarly, more details in a review can help readers reduce the uncertainty in evaluating product quality, increase the confidence in their purchase decisions, and help them make better decisions (Mudambi and Schuff 2010). Therefore, we propose the first hypothesis below.

H1. Review length has a positive effect on the helpfulness of the review.

According to ELM, situational factors can influence the likelihood of elaboration (Bhattacherjee and Sanford 2006; Petty and Cacioppo 1986). People are more likely to carefully scrutinize message content if the need to be accurate in judgment and decision-making is high, as such elaboration is the best way of achieving accuracy (Priester and Petty 1995). One key situational factor that can vary message recipients’ need to be accurate is source trustworthiness. As an essential dimension of source credibility, trustworthiness refers to the extent to which a source is believed to be honest and sincere rather than deceptive (Hovland et al. 1953).

In the online reviews context, we argue that reviewer trustworthiness can vary the likelihood of elaboration of review readers – the extent to which consumers will scrutinize the review content, thus moderating the impact of review length. A typical goal of review readers is to evaluate the accuracy of a number of usually contradictory reviewer opinions and to determine how much they could base their purchase decisions on particular reviews (Dabholkar 2006; Mudambi and Schuff 2010). When reviewers are trustworthy, their reviews are believed by readers to be more dependable and accurate. In this case, because the need for checking the accuracy of review content is low, consumers are more likely to accept the opinions and recommendation of reviewers with little thoughts rather than scrutinizing the review content. In contrast, reviews from untrustworthy reviewers result in heightened need for examining review accuracy, prompting review readers to scrutinize the review content in order to assess the validity of the reviewer options (see Priester and Petty 2003). In this situation, consumers tend to appreciate the additional details provided in a review to a greater extent due to their greater scrutiny of review content.
In sum, we propose that the positive effect of review length will be greater when the reviewer is less trustworthy.

**H2. The positive effect of review length on review helpfulness will be stronger for reviews of less trustworthy reviewers than reviews of more trustworthy reviewers.**

Another situational factor that can influence the likelihood of elaboration is distractions. People tend to spend time and effort in scrutinizing message content if they are not only motivated but also capable of evaluating the message (Petty and Cacioppo 1986). However, if such capability is constrained by distractions, people’s thinking and understanding of the message through the central route will be disrupted (Hovland et al. 1953), resulting in a lower likelihood of elaboration and a weaker effect of the central route on persuasion (Petty et al. 1976).

One source of distractions in online reviews is reading difficulty – the difficulty of comprehending a piece of text (Korfiatis et al. 2012). Reviews that are difficult to read should hinder the ability of readers to understand the review content and evaluate its quality. As reviews get more difficult to read, consumers will be less able to appreciate the additional details of a review and spend less effort in understanding the review content, thus weakening the positive effect of review length on review helpfulness. Thus, we propose the following hypothesis:

**H3. The positive effect of review length on review helpfulness will be stronger for easy-to-read reviews than hard-to-read reviews.**

### Interpersonal Influence of Emotional Intensity

In addition to review length, the expression of emotions in a review can also influence its helpfulness perceptions. However, the effect of emotions cannot be adequately explained by elaboration likelihood model, which does not clearly address the influence of emotions in persuasion (Nabi 1999). Instead, we develop the next set of hypotheses based upon the emerging literature on the social functions of emotions.

There has been an increasing recognition in recent decades that emotions not only are an intrapersonal phenomenon, but also serve important social functions (Parkinson 1996; Parkinson et al. 2004). An individual may recognize and perceive other people’s emotional expressions manifested through nonverbal and verbal manners (Atkinson et al. 2004; Barrett et al. 2007; Ekman and Friesen 1971; Lindquist et al. 2006). More importantly, the individual may draw inferences from others’ emotions to make sense of the situation (Russell 1991; Siemer 2008; Van Kleef 2010).

In online reviews, readers of emotional reviews may draw inferences from reviewers’ emotional expressions and inquire into the motives of reviewers. Emotions in our context refer to subjective feelings targeted at a specific product or purchase experience (Yin et al. 2014), whereas non-emotional content refers to objective evaluations of the pros and cons of the product. Furthermore, we define emotional intensity as the percentage of emotional content in a review following the prior literature (see Fujita et al. 1991; Kahn et al. 2007).

In particular, we argue that consumers may draw negative inferences from emotional expressions of reviewers, possibly due to concerns about online fraud and the manipulation of online reviews (Xiao and Benbasat 2011). Manipulative reviews are those that appear to be a truthful account of a real purchase experience but are, in fact, written by vendors associated with the product (Hu et al. 2011b). In other words, manipulative reviews represent a type of online fraud. Previous studies have found that product vendors tend to exaggerate the emotional cues in an attempt to create more viral messages and better persuade potential consumers (Banerjee and Chua 2014; Dobele et al. 2007). Linguistic analyses of compute-mediated communications also suggest that liars tend to use more emotional words than truth-tellers (Banerjee and Chua 2014). Reviewers’ attempt to win over the consumers through intensive use of emotions can give rise to suspicion of ulterior motives, leading to reduced credibility of the review (see Tang and Hall 1995). Although emotion-intensive reviews are not necessarily fake or fabricated reviews, our claim only requires that review readers have a tendency to associate such reviews with online fraud. As a result, review readers may associate excessive use of emotions with fabricated reviews, thus rating such reviews less helpful.

**H4. Emotional intensity of a review has a negative effect on the helpfulness of the review.**
In addition, we argue that the negative effect of emotional intensity on review helpfulness is more likely to apply to positive emotions than negative emotions. Because of the association between online reviews and product sales (Chevalier and Mayzlin 2006; Clemons et al. 2006), vendors have a strong incentive to manipulate reviews in a positive way in order to influence consumers’ purchase decisions (Hu et al. 2012). Empirical evidence from Amazon indicates that many product vendors are continuously monitoring online reviews and fabricating positive reviews to counteract the impact of negative reviews, and that review manipulation is more likely to occur for low quality products with low average ratings (Hu et al. 2011a; Hu et al. 2011b). Given the predominance of positive reviews available online compared with negative reviews (Pavlou and Dimoka 2006; Resnick and Zeckhauser 2002), customers are in general more suspicious of positive emotions than negative emotions. There was also evidence that liars use more positive emotions than truth-tellers, while the difference in the use of negative emotions does not reach significance (Zhou et al. 2004). Therefore, we predict that intensive emotions are more likely to be associated with manipulative behaviors for reviews with positive emotions than for reviews with negative emotions, and that the negative effect of emotional intensity will be stronger for the former than the latter.

**H5. The negative effect of the intensity of positive emotions on review helpfulness will be stronger (more negative) than the negative effect of the intensity of negative emotions.**

However, the negative effect of emotional intensity can be mitigated if the emotions are perceived sincere and authentic. Although intensive emotions can prompt readers to doubt the motives of the reviewer, emotions are inevitable and commonplace in any consumption experience (Hirschman and Holbrook 1982). The expression of emotions is also one of the major reasons why people write reviews (Berger 2011). Whether or not emotions will have a negative impact depends on their authenticity perceived by observers. According to the recent literature on social functions of emotions, inauthentic emotional expressions tend to be perceived as manipulative attempts to influence others (Côté et al. 2013). If emotions are perceived to be inauthentic, observers are more likely to react negatively (Hareli and Rafaeli 2008; Van Kleef et al. 2012).

We propose that reviewer trustworthiness is one of the factors that can directly influence perceived authenticity of emotions and mitigate consumers’ suspicion of review manipulation. The trustworthiness of reviewers signals the extent to which their reviews are truthful and dependable. Specifically, reviews from highly trustworthy reviewers are generally believed to be honest and sincere (Hu et al. 2008). As a result, it is reasonable to expect that intensive use of emotional words, especially the expression of positive emotions, is justified as truthful representations of reviewers’ real experiences if the reviews are from trustworthy sources. However, if the reviews are from less trustworthy sources, intensive expressions of emotions, especially positive emotions, are more likely to trigger review readers’ suspicion of the manipulative motives of reviewers. Therefore, we propose the following hypothesis.

**H6. The negative effect of a) emotional intensity and b) positive emotional intensity on review helpfulness will be stronger for reviews of less trustworthy reviewers than reviews of more trustworthy reviewers.**

Finally, we propose that reading difficulty can also affect the perceived authenticity of emotions based on the common understanding that fake reviews tend to be easier to read. The writing process of real reviews requires the recall and organization of pros and cons in a cognitively complicated manner. In general, vendors fabricating reviews cannot write in this way due to the cognitive taxing nature of review writing and a lack of real memories/experiences to base their reviews on (Zhou et al. 2004). In addition, fabricated reviews typically make use of words and sentences that are easier to understand in order to catch the attention of more consumers and influence a larger audience (Banerjee and Chua 2014). Linguistic analyses also find evidence that liars use words of lower cognitive complexity (Newman et al. 2003; Zhou et al. 2004). Thus, consumers are more likely to treat reviews as inauthentic if the reviews are easier to read, and we propose the following hypothesis. Figure 1 presents our theoretical framework.

**H7. The negative effect of a) emotional intensity and b) positive emotional intensity on review helpfulness will be stronger for easy-to-read reviews than hard-to-read reviews.**
Method

To test our hypotheses, we collected and analyzed actual reviews from Epinions.com, a popular third-party product review website. At Epinions.com, community members can share their opinions on products by writing reviews. More importantly, members can evaluate how helpful others’ reviews are along a scale ranging from “Very Helpful,” “Helpful,” “Somewhat Helpful,” to “Not Helpful.” Such a scale enables us to quantify the helpfulness of reviews – our major dependent variable – in a more precise manner than the commonly used ratio measure aggregated from “Yes/No” votes (e.g., Mudambi and Schuff 2010; Yin et al. 2014).

In addition, Epinions.com supports a mechanism that other review sites do not have, “web of trust.” This unique mechanism allows community members (i.e., trustors) to express their trust in other members (i.e., trustees who are trusted in the relationship) whose reviews and ratings they have consistently found to be valuable. The number of trustors is displayed prominently in the profile of each reviewer; as a reviewer is trusted by more members in “web of trust,” he/she should be regarded as more trustworthy (Ku et al. 2012). This mechanism makes it possible to operationalize reviewer trustworthiness, which is proposed as a moderator in a number of our hypotheses.

Data Collection

We collected the data on reviews from Epinions.com (Figure 2) in August 2013. We began by retrieving the reviews of all products in the 7 representative categories that have been studied in the prior literature, including digital camera, printer, calculator, cosmetics, video game, home theater system, and music CD. Previous studies have classified the first three product categories as search goods and the rest as experience goods (see Hung and Wyer 2009; Mudambi and Schuff 2010; Xiao and Benbasat 2007). For each review, we recorded the rating given to the focal product, review content, and helpfulness evaluation. We also collected data that can be used to derive the characteristics of reviewers who contributed the reviews, including the detailed “web of trust” of each reviewer (i.e., who trust the focal review and when) and reviews written by each reviewer. Finally, we collected data on product characteristics, such as product average ratings and prices. In total, we collected 7937 product reviews.
Variables

The dependent variable is the review helpfulness (Helpfulness). Review readers can rate a product review as “Very Helpful,” “Helpful,” “Somewhat Helpful,” or “Not Helpful.” For each product review, Epinions.com aggregates all the ratings and presents an overall rating of helpfulness for the review. Following Ku et al. (2012), we treated review helpfulness as a Likert-scale measure, and assigned the value of 3 to a “Very Helpful” review, 2 to a “Helpful” review, 1 to a “Somewhat Helpful” review, and -1 to a “Not Helpful” review.

The major independent variables of interest are the length of a review (Review Length) and the intensity of emotional words in a review (Emotional Intensity). First, we measured review length by counting the number of words in a review. Second, emotional intensity was computed by Linguistic Inquiry and Word Count (LIWC) software developed by Pennebaker and his colleagues (2007). The reliability and validity of LIWC has been demonstrated by its developers (Pennebaker et al. 2007; Pennebaker and Francis 1996), and this software has been widely used to efficiently quantify psychological constructs in diverse disciplines, such as psychology, linguistics, marketing, and information systems (Bantum and Owen 2009; Berger and Milkman 2012; Tausczik and Pennebaker 2010; Yin et al. 2014). Therefore, we deem LIWC an appropriate tool to measure our emotion-related variables. Specifically, LIWC calculates emotional intensity by dividing the number of emotional words identified in its dictionary by the number of total words in a review. LIWC can further break it down to the percentage of positive emotional words (Positive Emotional Intensity) and the percentage of negative emotional words in a review (Negative Emotional Intensity).

We also measured the two moderators in this study. First, given a focal review, the trustworthiness of its author was measured by the number of trustors the reviewer has up to the date the review was released (Reviewer Trustworthiness). Second, following the previous research (Korfiatis et al. 2012; Yin et al. 2014), we quantify the difficulty of reading a review (Reading Difficulty) by computing the Coleman-Liau Index, an estimate of the U.S. grade level that a reader must have to understand the text. A higher value of the index indicates more difficulty in comprehending the text.

We controlled for a series of relevant variables in the model. First, we included the number of reviews a reviewer has posted before the focal review (Reviewer Experience). Second, we controlled for the product rating (from 1 star to 5 star) given in the review (Product Rating) and the number of days since the review was posted (Days). Third, we controlled for the total number of helpful votes casted by review readers (#Total Votes). Fourth, we included a product type dummy (Product Type) to account for systematic differences across product types. This variable was coded as 0 for search goods and 1 for experience goods. Finally, we controlled for other product-level variables, including the average rating of a product (Average Product Rating) and the average price of a product (Average Product Price).
Analysis and Results

We used hierarchical OLS regression analysis to test our hypotheses. We first entered the control variables, then added the independent variables and the moderators, and finally added the interaction effects. As suggested by Aiken and West (1991), we mean-centered the independent variables and moderators to mitigate multicollinearity and to facilitate the interpretation of coefficients of interaction terms. Table 1 reports the means, standard deviations, and correlations of the variables in the analysis.

Hypothesis 1 posited that a review with a greater number of words is rated more helpful. In model 2 of Table 2, Review Length was positively associated with Helpfulness ($\beta = 0.0007$, $p < .01$), providing evidence for Hypothesis 1. This result suggests that, if the length of a review increases by 1400 words, its helpfulness will increase by almost one point (e.g., from “Helpful” rating to “Very Helpful” rating).

Table 1: Means, Standard Deviations, and Correlations

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Helpfulness</td>
<td>1.780</td>
<td>1.253</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Review Length</td>
<td>557.466</td>
<td>617.356</td>
<td>0.378</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Emotional Intensity</td>
<td>5.088</td>
<td>2.358</td>
<td>-0.117</td>
<td>-0.179</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Positive Emotional Intensity</td>
<td>3.751</td>
<td>2.071</td>
<td>-0.140</td>
<td>-0.192</td>
<td>0.873</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Negative Emotional Intensity</td>
<td>1.316</td>
<td>1.144</td>
<td>0.014</td>
<td>-0.020</td>
<td>0.477</td>
<td>-0.010</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>6. Reviewer Trustworthiness</td>
<td>20.470</td>
<td>84.255</td>
<td>0.225</td>
<td>0.441</td>
<td>-0.111</td>
<td>-0.094</td>
<td>-0.060</td>
<td>1.000</td>
</tr>
<tr>
<td>7. Reading Difficulty</td>
<td>9.091</td>
<td>1.482</td>
<td>0.150</td>
<td>0.174</td>
<td>0.006</td>
<td>-0.050</td>
<td>0.101</td>
<td>0.131</td>
</tr>
<tr>
<td>8. Reviewer Experience</td>
<td>0.538</td>
<td>2.722</td>
<td>0.154</td>
<td>0.416</td>
<td>-0.084</td>
<td>-0.074</td>
<td>-0.039</td>
<td>0.602</td>
</tr>
<tr>
<td>9. Product Rating</td>
<td>4.051</td>
<td>1.84</td>
<td>0.097</td>
<td>0.049</td>
<td>0.060</td>
<td>0.182</td>
<td>-0.206</td>
<td>0.013</td>
</tr>
<tr>
<td>10. Days</td>
<td>3196.518</td>
<td>1470.893</td>
<td>0.236</td>
<td>-0.221</td>
<td>0.202</td>
<td>0.143</td>
<td>0.158</td>
<td>-0.175</td>
</tr>
<tr>
<td>11. #Total Votes</td>
<td>12.826</td>
<td>16.711</td>
<td>0.293</td>
<td>0.488</td>
<td>-0.123</td>
<td>-0.116</td>
<td>-0.043</td>
<td>0.599</td>
</tr>
<tr>
<td>12. Product Type</td>
<td>0.414</td>
<td>0.493</td>
<td>0.090</td>
<td>-0.059</td>
<td>0.387</td>
<td>0.291</td>
<td>0.271</td>
<td>-0.062</td>
</tr>
<tr>
<td>13. Average Product Rating</td>
<td>3.935</td>
<td>0.835</td>
<td>0.053</td>
<td>-0.013</td>
<td>0.154</td>
<td>0.181</td>
<td>-0.008</td>
<td>-0.008</td>
</tr>
<tr>
<td>14. Average Product Price</td>
<td>237.197</td>
<td>767.116</td>
<td>-0.017</td>
<td>0.024</td>
<td>-0.042</td>
<td>-0.013</td>
<td>-0.062</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Reading Difficulty</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Reviewer Experience</td>
<td>0.130</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Product Rating</td>
<td>0.045</td>
<td>0.023</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Days</td>
<td>-0.029</td>
<td>-0.134</td>
<td>0.048</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. #Total Votes</td>
<td>0.143</td>
<td>0.316</td>
<td>0.024</td>
<td>-0.098</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Product Type</td>
<td>0.074</td>
<td>-0.062</td>
<td>0.071</td>
<td>0.359</td>
<td>-0.051</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Average Product Rating</td>
<td>0.081</td>
<td>0.009</td>
<td>0.412</td>
<td>0.132</td>
<td>-0.002</td>
<td>0.226</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>14. Average Product Price</td>
<td>0.000</td>
<td>0.009</td>
<td>0.005</td>
<td>-0.059</td>
<td>-0.001</td>
<td>-0.109</td>
<td>-0.012</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: Review Length, Emotional Intensity, Positive Emotional Intensity, Negative Emotional Intensity, Reviewer Trustworthiness and Reading Difficulty are not mean-centered in Table 1.
<table>
<thead>
<tr>
<th>Table 2: OLS Regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Product Type</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Average Product Price</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Average Product Rating</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong># Total Votes</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Days</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Product Rating</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reviewer Experience</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reading Difficulty</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reviewer Trustworthiness</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Review Length</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Emotional Intensity</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reviewer Trustworthiness × Review Length</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reviewer Trustworthiness × Emotional Intensity</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reading Difficulty × Review Length</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reading Difficulty × Emotional Intensity</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Positive Emotional Intensity</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Negative Emotional Intensity</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Hypothesis 2 suggested that the positive relationship between review length and review helpfulness is moderated by reviewer trustworthiness. In model 3 of Table 2, the interaction effect of **Reviewer Trustworthiness** and **Review Length** was significant ($\beta = -0.00002$, $p < .01$). To illustrate the nature of this significant interaction effect, we followed the recommended procedure by Aiken and West (1991), and plotted the predicted values of the dependent variable at one standard deviation above and one standard deviation below the mean of the independent variable. Figure 3a illustrates the interaction between **Reviewer Trustworthiness** and **Review Length**. The marginal effect of review length for high trustworthy reviewers ($\beta_{\text{High Trustworthy Reviewers}} = .0007$, $p < .01$) is less positive than that for low trustworthy reviewers ($\beta_{\text{Low Trustworthy Reviewers}} = .0010$, $p < .01$), leading support to Hypothesis 2. These results indicate that the positive impact of review length is significantly greater for reviewers who are low in trustworthiness, and that low trustworthy reviewers benefit more by expending more effort in the review writing process.

Hypothesis 3 suggested that the positive relationship between review length and review helpfulness is also moderated by the difficulty of reading a review. In model 3 of Table 2, the interaction effect of **Reading Difficulty** and **Review Length** was significant ($\beta = -0.0001$, $p < .01$). Figure 3b illustrates this interaction effect and shows that the marginal effect of review length for hard-to-read reviews ($\beta_{\text{Hard-to-Read}} = .0007$, $p < .01$) is less positive than that for easy-to-read reviews ($\beta_{\text{Easy-to-Read}} = .0009$, $p < .01$). This result suggests that the positive effect of review length will be weaker when reviews are hard-to-read than when reviews are easy-to-read. Therefore, Hypothesis 3 was supported.

Hypothesis 4 proposed that a review with a higher intensity of emotional words would be rated less helpful. In model 2 of Table 2, **Emotional Intensity** was negatively related to **Helpfulness** ($\beta = -0.0651$, $p < .01$), leading support to Hypothesis 4. This result suggests that a 15% increase in the intensity of emotion words in a review leads to a decrease of more than one point in review helpfulness.
Hypothesis 5 postulated that positive emotional intensity has a stronger effect on review helpfulness than negative emotional intensity. In model 4 of Table 2, the estimated coefficients of positive and negative emotional intensity are -.0747 and -.0298, respectively. The difference between these two coefficients was negative and significant ($F$-value = 4.98, $p < .05$), so Hypothesis 5 was supported. Our results suggest that
a one-point decrease in review helpfulness can be associated with a 13% increase in the intensity of positive emotional words, while the association with the intensity of negative emotional words is not significant.

Hypothesis 6a suggested that the negative relationship between emotional intensity and review helpfulness depends on reviewer trustworthiness. In model 3 of Table 2, the interaction effect of Reviewer Trustworthiness and Emotional Intensity was significant ($\beta = .0004, p < .01$). Figure 3c plots this significant interaction and shows that the marginal effect of emotional intensity for reviewers with a low level of trustworthiness was negative and significant ($\beta_{\text{Low Trustworthy Reviewers}} = -.0806, p < .01$), while the marginal effect for reviewers with a high level of trustworthiness was not significant at the .05 significance level ($\beta_{\text{High Trustworthy Reviewers}} = -.0203, p < .10$). These results suggest that the negative impact of emotional intensity on review helpfulness goes away for reviewers who are trustworthy. Thus, Hypothesis 6a was supported.

Hypothesis 6b proposed that the effect of positive emotional intensity on review helpfulness depends on reviewer trustworthiness. In the model 5 of Table 2, the interaction effect of Reviewer Trustworthiness and Positive Emotional Intensity was significant ($\beta = .0004, p < .01$). Figure 3d illustrates this significant interaction and shows that the marginal effect of positive emotional intensity for low trustworthy reviewers was negative and significant ($\beta_{\text{Low Trustworthy Reviewers}} = -.0900, p < .01$), while the marginal effect for high trustworthy reviewers was not significant ($\beta_{\text{High Trustworthy Reviewers}} = -.0209, p > .10$). This suggests that the moderation effect of reviewer trustworthiness observed above applies to positive emotions. Therefore, Hypothesis 6b was supported.

Hypothesis 7a suggested that the negative relationship between emotional intensity and review helpfulness is moderated by the difficulty of reading a review. In model 3 of Table 2, the interaction effect of Reading Difficulty and Emotional Intensity was significant ($\beta = .0116, p < .05$). Figure 3e plots this significant interaction and shows that the marginal effect of emotional intensity for easy-to-read reviews was negative and significant ($\beta_{\text{Easy-to-Read}} = -.0333, p < .05$), while the marginal effect for hard-to-read reviews was not significant ($\beta_{\text{Hard-to-Read}} = -.0270, p > .10$). These results indicate that the negative effect of emotional intensity goes away for reviews that are hard to read. Thus, Hypothesis 7a was supported.

Hypothesis 7b postulated that the negative effect of positive emotional intensity is weaker as reading difficulty increases. In the model 5 of Table 2, the interaction effect of Reading Difficulty and Positive Emotional Intensity did not reach significance at the .05 significance level ($\beta = .0109, p < .10$). Therefore, Hypothesis 7b was not supported.

**Discussion and Conclusion**

In this study, we find that the common understanding of “the more the better” applies to the length of a review, but not to the emotional content expressed by reviewers. To the contrary, we find that the intensity of emotions, especially positive emotions, expressed in a review is negatively associated with review helpfulness. These results are consistent with the notion that a higher level of emotional expressions may prompt readers to make negative attributions and associate such reviews with review manipulation and online fraud (Banerjee and Chua 2014; Dobele et al. 2007). Therefore, reviewers are less likely to craft a helpful review if they express their positive feelings without constraints.

In addition, our results suggest that the respective effects of review length and emotional intensity can be attenuated by reviewer and content characteristics. Specifically, we find that the positive effect of review length is attenuated when a review comes from a more trustworthy reviewer or when the review is written in a sophisticated manner. Under such occasions, the detrimental impact of positive emotions is also likely to go away because emotional reviews are more likely to be considered authentic and associated with real customer experiences than online fraud.

**Theoretical Implications**

Our paper makes several unique contributions to the literature on online reviews and user-generated content. First, we contribute to the previous research on review helpfulness by exploring emotion-related antecedents. While prior studies have demonstrated repeatedly the effect of cognitive factors, such as review length and ratings (e.g., Baek et al. 2012; Forman et al. 2008; Mudambi and Schuff 2010), little
research has addressed how a review's emotional content influences its helpfulness evaluations. Our findings suggest that the intensity of emotional content matters, and that positive emotions have a greater effect on review helpfulness than negative emotions. Furthermore, we find that review length and emotional intensity have opposite impacts on review helpfulness. That is, we demonstrate that “the more the better” is not a universal principle guiding helpfulness evaluation of reviews; it only applies to the length of a review, not to the intensity of emotions. These findings challenge the conventional wisdom of “the more the better” and emphasize the importance of taking into account emotional factors in studying online consumer behavior.

Second, our study is among the first attempts to examine the interpersonal effect of emotions in consumer decision-making (Yin et al. 2014). Emotions were traditionally studied as an intrapersonal phenomenon, but they can also serve important social functions (Parkinson 1996; Parkinson et al. 2004). The interpersonal impact of emotions is especially pervasive and long-lasting in online reviews, whereby the feelings expressed by reviewers can influence the purchase decisions of thousands of consumers who consult the reviews. In particular, we propose an explanation for the detrimental effect of positive emotions – concerns about review manipulation and online fraud, and provide empirical evidence consistent with this explanation. Therefore, this paper deepens our understanding of how emotions are interpreted and attributed in interpersonal settings, and extends the examination of the social functions of emotions to online word-of-mouth.

Third, this paper enriches our understanding of the boundary conditions for the effects of review length and emotional intensity on review helpfulness. Specifically, we extend the literature on elaboration likelihood model by investigating the moderating roles of source and content variables. Prior studies have investigated the moderating roles of certain reviewer characteristics (e.g., reviewer expertise) (e.g., Li et al. 2013; Ma et al. 2013), but little research has paid attention to reviewer trustworthiness, another essential dimension of source credibility. In addition, the difficulty of reading a review has received growing interest as an independent variable in the context of online review (Korfiatis et al. 2012), but its moderating effect in consumer decision-making has not been fully acknowledged. In addition, we find that these factors also moderate the detrimental effect of emotional intensity: increasing the intensity of emotions does no harm to review helpfulness if the reviewer is trustworthy or if the review content is written in a sophisticated manner. These findings are consistent with the argument that perceived authenticity of emotions influences the extent to which observers will make negative inferences (Harel and Rafaeli 2008; Van Kleef et al. 2012).

**Practical Implications**

Our findings also have the potential to offer important practical implications for online review contributors, review platforms, and product vendors. For review contributors who strive to craft the most helpful reviews, our findings provide some promising guidelines. First, consistent with prior studies, the easiest way is to increase the amount of information. This guideline is especially critical for new reviewers who have not established sufficient trustworthiness from their communities. In addition, our findings suggest that readability matters, and that additional information contained in a review has a greater impact on review helpfulness if the information is easier to read.

Second, our study suggests that reviewers need to be aware of the use of emotional words. Review readers may perceive a review less helpful if the intensity of emotions is higher. However, this negative effect of emotional intensity only applies to positive emotions, not to negative emotions. This finding suggests that reviewers who are unsatisfied with a product or service need not worry too much about the expression of their negative feelings. On the other hand, reviewers who favor a product or service should pay more attention to their expressions of positive emotions. Our study further suggests that one way to mitigate this negative effect of positive emotions is to build a profile that can signal the trustworthiness of reviewers.

Our paper also provides some guidelines for the design of review websites. For example, websites will be better off implementing features concerning the content of a review to encourage high quality reviews. Our findings suggest that review length has a positive effect on review helpfulness, but emotional intensity has a negative effect. Thus, review websites could design a feature to notify review contributors when a review contains not enough information or too many emotional words. In addition, our study demonstrates the importance of reviewer trustworthiness on a review website. We find that the negative
effect of emotional intensity goes away if a review comes from a highly trustworthy reviewer. This finding implies that the prominent display of trustworthiness cues in reviewer profiles is very critical on a review website.

Finally, this study is useful for product vendors who want to identify helpful reviews for marketing purposes. In general, if customers can understand the benefits of one product earlier than the benefits of alternative products, the vendor of the former product is more likely to win over customers. However, the identification of the most helpful and persuasive reviews is especially challenging when a product is new on the market or when the product reviews have received very few helpfulness votes. Our study suggests that a helpful review can be evaluated from four aspects: review length, emotional intensity, reviewer trustworthiness, and reading difficulty.

Conclusion

Review websites for products and services are increasing important for consumer decision-making. Previous studies have focused on the amount of information contained in a review. Our study shows that, in addition to review length, emotional intensity is also significantly related to review helpfulness, but they have opposite effects. Moreover, these relationships are moderated by source and content characteristics—reviewer trustworthiness and review reading difficulty. Overall, this work enhances current understanding of how to craft a helpful review.

References