In Search of Negativity Bias:
An Empirical Study of Perceived Helpfulness of Online Reviews

Philip Fei Wu
Surrey Business School
University of Surrey
Guildford, United Kingdom
Email: f.wu@surrey.ac.uk
Phone: +44(0)1483 68 6327

In Search of Negativity Bias:
An Empirical Study of Perceived Helpfulness of Online Reviews

Abstract

A basic tenet of psychology is that the psychological effects of negative information outweigh those of positive information. Three empirical studies show that the negativity bias can be attenuated or even reversed in the context of electronic word-of-mouth (eWoM). The first study analyzes a large sample of customer reviews collected from Amazon.com and concludes that negative reviews are no more helpful than positive ones when controlling for review quality The second study follows up with a virtual experiment that confirms the lack of negativity bias in evaluating the helpfulness of online reviews. The third study demonstrates that the negativity effect can be reversed by manipulating the baseline valences. This work challenges the conventional wisdom of "bad is stronger than good" and contributes to the understanding of the eWoM phenomenon.

People often seek the opinions of others to assist their decision making (Soll & Larrick, 2009). There is a large body of research documenting the influences of critics or friends on consumers' product evaluations and choices (Arndt, 1967; Gilly, Graham, Wolfinbarger, & Yale, 1998; Liu, 2006; S.-B. Park & Park, 2013). Continuation of this research theme in the Internet era has led to studies of consumer-generated online product reviews (Chevalier & Mayzlin, 2006; Dellarocas, 2003; Duan, Gu, & Whinston, 2008a). Many e-commerce companies such as Amazon and eBay, implement online review systems that solicit and publish customers' opinions about the products they have purchased. Hennig-Thurau and Walsh (2004) refer to this type of online feedback as "electronic word-of-mouth" (eWoM). Due to the openness and hyperconnectivity of the Internet, eWoM is being generated at an unprecedented scale and speed (Dellarocas, 2003).

However, work in this area is relatively fragmented and the empirical findings are sometimes inconclusive or conflicting (Dellarocas, 2003; Hu, Liu, & Zhang, 2008). This is due to a variety of reasons including the complexity of consumer decision marking (D.-H. Park, Lee, & Han, 2007; de Valck, van Bruggen, & Wierenga, 2009), the lack of accurate sales data (Chevalier & Mayzlin, 2006), the difficulty of performing qualitative analyses with large text corpora (Ghose & Ipeirotis, 2009), and the variations in products, consumers, and online shopping contexts (Forman, Ghose, & Wiesenfeld, 2008; Huang, Lurie, & Mitra, 2009; Weathers, Sharma, & Wood, 2007). In light of the challenges involved in establishing a direct link between eWoM and sales, researchers have turned their attention to the perceived value of online reviews for the consumer. An important and interesting research question in this context is whether online

consumers perceive negative reviews to be more helpful than positive reviews. Such "negativity bias" would have immediate consequences for online marketing managers and researchers.

Situated in a rich tradition of psychological research on information valence, this research explores the relationship between information valence and the perceived helpfulness of online reviews. Unlike previous academic work in this area, the present research takes a mixed-method approach and combines data mining and virtual experimentation to establish a more solid understanding. The first empirical study is aimed at reproducing the negativity bias by analyzing a large dataset of Amazon book reviews. The second and the third studies follow up with virtual experiments to extend the research from detecting to explaining the negativity effect.

The rest of the paper is organized as follows. A brief review of existing research on eWoM and online reviews is followed by an explanation of the negativity bias and its relevance for this research. The three empirical studies and the results of the data analysis are presented in detail. Finally, the contributions and practical implications of the research are discussed.

RESEARCH BACKGROUND

Prior Research on eWoM

A consistent theme in marketing research on word-of-mouth is the effect of negative word-of-mouth (Berger, Sorensen, & Rasmussen, 2010; Herr, Kardes, & Kim, 1991; Richins, 1983). Although it might seem straightforward to reason that negative word-of-mouth will hurt product

sales and brand evaluation, there are conflicting findings with regard to the association of negative eWoM and sales. For example, Basuroy, Chatterjee, and Ravid (2003) find that unfavorable reviews from film critics reduce box office revenues, a finding supported by Dellarocas, Zhang, and Awad's (2007) analysis of user reviews posted on *Yahoo! Movies* discussion boards. Surprisingly, using the same *Yahoo! Movie* data source, Liu (2006) and Duan, et al. (2008b) contend that it is the volume of user postings rather than the valence of reviews, that has a significant impact on movie box office revenues. There is a similar contradiction in studies of Amazon book reviews: while Chevalier and Mayzlin (2006) show that one-star reviews on Amazon.com hurt book sales, Forman et al. (2008) find no significant relationship between review valence and sales.

In light of the challenges involved in establishing a solid link between eWoM and sales, recent research has begun to investigate the significance of eWoM from the viewpoint of the consumer. Indeed, if one expects online reviews to influence consumer attitudes to products, which in turn lead to purchase decisions, it is crucial to understand how the consumer considers reviews in the process of making a purchase decision. Researchers have begun to ask the question "What do consumers think of online reviews and why?", and also "How do online reviews influence consumers?" For example, both Mudambi and Schuff (2010) and Ghose and Ipeirotis (2009) refer to the qualitative characteristics of online reviews, such as review depth and subjectivity, to explore which kinds of reviews are perceived by consumers to be more helpful. Forman et al. (2008) suggest that consumers use reviewer identity information in electronic markets to supplement product information when using online reviews to help their decision making. By comparing product reviews on four national Amazon sites (U.S., U.K.,

German, and Japanese), Danescu-Niculescu-Mizil, Kossinets, Kleinberg, and Lee (2009) note

national differences among reviews in terms of review variance and review helpfulness.

These studies rely on mining review data from e-commerce or online community websites.

A typical approach is to program a robot crawler to automatically crawl the target website and

extract relevant data such as user reviews and ratings (e.g., Duan et al., 2008a). In recent years,

some e-commerce sites have been implementing a feature that allows online visitors to rate the

"helpfulness" of consumer-generated product reviews. This interesting feature produces

aggregated numbers of helpfulness votes which are used as a proxy measure for the perceived

value of online reviews by most academic publications (Danescu-Niculescu-Mizil et al., 2009;

Mudambi & Schuff, 2010; Wu, van der Heijden, & Korfiatis, 2011). This data mining approach

yields many interesting findings but has some major methodological limitations. Firstly, only a

small portion of review readers cast helpfulness votes, and there is no practical method to

remedy non-respondent bias. Secondly, there are many unobservable or uncontrollable variables

related to reviewers (e.g., motive), review readers (e.g., personality), and review texts (e.g.

relevant information bits) which could influence the perceived helpfulness of a particular review.

While qualitative analysis of review content is possible, few have attempted this because such

data analysis is deemed expensive.

The Negativity Bias

A basic tenet of psychology is that the psychological effects of negative information

outweigh those of positive information. By and large, people pay more attention to bad than

good news, and they take criticism more seriously than praise. While this "positive-negative asymmetry" (Peeters, 1971; Taylor, 1991), or "negativity bias" (Ito, Larsen, Smith, & Cacioppo, 1998; Rozin & Royzman, 2001) has been confirmed repeatedly by researchers in many social science disciplines, the premise enjoys prominent attention in the impression formation literature. Numerous studies in this domain confirm a perceptual bias in which negative information about a stimulus person carries more weight and has a bigger effect on impression, than positive information (Rozin & Royzman, 2001). An early study by Riskey and Birnbaum (1974) concluded that a person's overall impression is determined mostly by his bad as opposed to his good deeds. For instance, a sex scandal can ruin a politician's image no matter how many his years of good marriage. A plausible explanation of the negativity bias in impression formation is that negative behaviors contain more distinctive information than positive behaviors, and therefore are perceived as more diagnostic in categorizing individuals (Skowronski & Carlston, 1989). For example, the behavior of cheating reveals more about a person's honesty than the behavior of truth telling. These observations in impression formation accord with a wider theoretical account of negativity bias which refers to a general bias in humans to give greater weight to negative entities. For example, Taylor (1991) studies the negativity bias from a physiological perspective and finds that there is more physiological arousal from bad events, and that negative stressors have a more powerful effect on health than equivalent positive stressors.

While mainstream psychology research on negativity bias centers around personal traits and moral judgments (i.e., "bad guy" versus "good guy"), Amabile and colleague (Amabile & Glazebrook, 1982; Amabile, 1983) propose an interpersonal evaluation theory that extends the

negativity bias hypothesis to evaluate people's intelligence and knowledgeability. The theory suggests that an evaluator's negative assessment of a stimulus object is likely to promote a favorable impression of the evaluator's intelligence. The effect is strengthened if the assessment is well-reasoned and elaborated at some length (Amabile, 1983). In her experiment, Amabile asked participants to read two book reviews demonstrating extremes of valence (one extremely positive and the other extremely negative) and rate the perceived intelligence of the reviewers. The results showed that the negative reviewer was perceived as brighter than the positive reviewer. Interpreting the results in an impression formation context, Amabile argues that negative criticism can be used as an impression management strategy since negatively critical evaluations impress observers as being more intelligent.

From moral judgment to intelligence evaluation, psychologists are interested more in the perceptual bias toward the human evaluators than the information objects. Does a similar negativity bias exist in people's impressions of stimulus information objects such as consumer product reviews? Are negatively valenced product reviews perceived to be more useful? Prior research in marketing shows that consumers tend to search for negative word-of-mouth in situations where they lack information and experience (Herr et al., 1991). Linking to Amabile's (1983) book review experiments discussed above, the alleged intelligence incorporated in negative comments implies new information that may help reduce uncertainty in the consumer's decision making (Dowling & Staelin, 1994; Kim, Ferrin, & Rao, 2008). Therefore, it seems reasonable to assume that consumers are likely to perceive negatively valenced product reviews as more helpful than positively valenced reviews. Three empirical studies were conducted to explore this negativity bias hypothesis.

EMPIRICAL STUDIES

Study 1

Apparently, the valence of reviews is not the only factor that determines the perceived helpfulness of reviews. The potential value of a review materializes only if the information contained in the review is credible and accessible to a general audience. A well-written review is likely to contribute to the favorable perception of helpfulness because it reduces the reader's cognitive effort in information consumption and at the same time, increases the credibility of the review. Research into information presentation, for example, has long demonstrated that the delivery of information, such as clarity, and detail of writing, has a significant impact on the reader's perception of its credibility (Fogg et al., 2003; Metzger, 2007). In developing interpersonal evaluation theory, Amabile (1983) also considers the quality dimension of evaluations and states that the alleged intelligence of a negative evaluator becomes more credible if the negative judgment is elaborated at some length. In the context of online consumer reviews, a well-written and substantive review is likely to provide more product details, in a more convincing way. A more recent study by Ghose and Ipeirotis (2009) finds that the readability and linguistic correctness of Amazon reviews is associated with votes about the helpfulness of the reviews: an increase in the readability of reviews has a positive impact on perceived helpfulness whilst an increase in the proportion of spelling errors has a negative impact on helpfulness. Mudambi and Schuff (2010) examine the aspect of review depth (measured by the proxy variable "word count") and find that longer reviews generally are perceived as more helpful.

This first empirical study aims to investigate the extent to which the qualitative characteristics of reviews moderate the effect of negativity bias in evaluating the helpfulness of reviews. Following common practice related to studying online reviews, the researcher collected customer reviews of the top 100 best selling books from Amazon.com. The automated data crawling yielded a collection of 44,328 book reviews. After removing duplicate text reviews (e.g., hardcover and paperback versions of the same book) and video reviews, the final dataset for the analysis contained 40,755 customer reviews for 88 distinct book items. For each review, the numerical valence data (on the 1-5 star scale Amazon provides), the helpfulness vote, and the review text were recorded in the dataset. Two qualitative constructs are measured, based on analyzing review text: readability and length. The length of review is the number of words contained in each review and the readability of the review is measured by the Flesch Reading Ease (FRE) – a popular readability index designed to measure easiness of comprehension of a piece of text in standard English (Flesch, 1948; Kincaid, Fishburne, Rogers, & Chissom, 1975). The FRE scores are subject to an interval censoring technique and range from 0 to 100, with a higher score indicating easier reading. As a rule of thumb, a text with an FRE score of 0-30 is considered very difficult, and a score of 60-70 indicates the appropriate level of readability for the general public. The helpfulness of the review is quantified using a feature on Amazon provided at the bottom of each review that allows readers to evaluate the review by indicating "Yes" or "No" to the question: "Was this review helpful to you?" The results of this voting appear at the top of each review in the form of "[# of Yes votes] out of [# of all votes] found the following review helpful" (see Appendix 1 for an example). Thus, the share of evaluators who found the review to be helpful is used as proxy for helpfulness.

Table 1 provides the descriptive statistics for the book review sample. The average customer rating of the books is positively valenced (M = 4.15, SD = 1.13). Evaluation of the helpfulness of reviews also tends to be positive, with an average of 64% voters find a particular review helpful. The length of the reviews varies greatly from a single word to 5,658 words. However, most reviews are less than 200 words (M =155.51, SD =168.62). The average FRE score after applying an interval censoring procedure (0-100), is 67.35, which suggests that the reviews are of standard readability and are appropriate for general adult readers.

[Insert Table 1 about here]

To detect negativity bias in rating the helpfulness of reviews and the potential attenuating effects of review length and readability, two models were analyzed:

Model 1: *Helpfulness = α + β₁ Valence + β₂ Valence² + ε*

Model 1: $Helpfulness = \alpha + \beta_1\ Valence + \beta_2\ Valence^2 + \varepsilon$

Model 2: $Helpfulness = \alpha + + \beta_1\ Valence + \beta_2\ Valence^2 + \beta_3\ WordCount + \beta_4\ FRE + \varepsilon$

In the two models, *Helpfulness* of a review is operationalized as the ratio of "helpful" votes to total votes received for the review. *Valence* is the overall valence of the book review, quantified by Amazon's 5-star rating scale. Many online customer reviews, however, are neither purely negative nor purely positive. In the context of product reviews on Amazon, reviews with 2, 3, or 4 stars usually indicate a middle-ground and mixed attitudes. From the perspective of information diagnosticity, reviews with mixed attitudes are likely to provide more balanced evaluations of the books, and offer readers more diagnostic information (Eisend, 2013). For example, Mudambi and Schuff (2010) find that Amazon reviews with extremely high or

extremely low star ratings are associated with lower levels of helpfulness than reviews with moderate ratings. Hence, the models include a quadratic term *Valence²* to estimate a curvilinear concave-shaped relationship between rating and helpfulness.

The models were analyzed using both ordinary linear square (OLS) and Tobit regressions to cross check the robustness of each method. The two methods produced similar results for the coefficient estimates and level of significance. The OLS regression results are presented in Tables 2 and 3.

[Insert Table 2 about here]

[Insert Table 3 about here]

In Model 1, the positive coefficient of *Valence* ($\beta_1 = 0.288$, $p < 0.001$) and the negative coefficient of *Valence²* ($\beta_2 = -0.188$, $p < 0.001$) suggest a concave, curvilinear shaped relationship between the book rating and the perceived helpfulness of the book review: highly positive reviews are less likely to be voted helpful than reviews that contain some level of negativity, although there is an overall positive association between valence and helpfulness. This ambiguity is resolved in the Model 2 analysis, where the downward curvature observed in Model 1 disappears after factoring in the qualitative characteristics of the reviews. The quadratic term *Valence²* loses its statistical significance ($p = 0.146$) while *Valence* remains significant ($p < 0.01$). In other words, the expected negativity bias in perceiving review helpfulness is not present after controlling for word count and readability of the review text. The stepwise regression also

12

shows that Model 2 is an improvement over Model 1 with a significant increase in *F* value and overall model fit.

Study 2

While Study 1 provides initial evidence that the negativity bias hypothesis might not be applicable to the eWoM context, the data mining approach has the same methodological limitations discussed earlier. In order to control for intervening variables such as reviewer reputation and review wording variations, the researcher adopted a novel, virtual experimentation method in this follow-up study. The design of the experiment followed a recent study on restaurant reviews (Naylor, Lamberton, & Norton, 2011) where participants were exposed to artificial reviews and their attitudes and intentions measured under different manipulation conditions. Participants were recruited from Amazon Mechanical Turk (www.mtruk.com) or MTurk, the largest online labor marketplace where registered users perform small tasks for micro payments. Research shows that workers on MTurk are more demographically diverse and closer to the U.S. population as a whole than traditional college subject pools (Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010). MTurk workers can register as individuals or "requesters" who create and post tasks (or HITs, acronym for Human Intelligence Tasks), or as "workers" who complete the tasks in exchange for a payment. Previous studies demonstrate the validity and many benefits of running virtual experiments with MTurk workers (e.g., Horton, Rand, & Zeckhauser, 2011; Paolacci et al., 2010).

The researcher browsed dozens of the most rated restaurants on the popular travel review website TripAdvisor.com, and skimmed hundreds of customer reviews of these restaurants. The exercise suggested that although the reviews differed in valence and detail, two criteria were universal for evaluating dining experiences: quality of the food, and standard of service. Based on this observation, the researcher created three artificial reviews with different valence levels (extremely positive, mixed, and extremely negative) of a fictitious restaurant, XYZ Kitchen in New York City. Most of the sentences in the reviews were adapted from actual TripAdvisor reviews commenting on food quality and standards of service. Note especially that the researcher deliberately held constant the amount of information concerning the two aspects across all three reviews by using similar wording and word count (see Appendix 2). Thus, each participant received the same amount of information, about the same aspects of the restaurant. This helps to rule out the possibility that differences in perceiving the reviews are due to exposure to different amounts or types of information. To ensure that the restaurant reviews were properly valenced, the three reviews were pre-tested with 95 undergraduate management students. Each student read all the reviews and then gave a valence rating using a 5-point scale (1 = "extremely negative" and 5 = "extremely positive"). The ratings were then compared using location tests to validate the corresponding valence levels assumed by the researcher. The results showed that all t values were small and p values were greater than 0.05: $M_{positive} = 4.99$, $t(94) = -1.00$, $p = 0.32$; $M_{mixed} = 3.02$, $t(94) = 0.82$, $p = 0.42$; $M_{negative} = 1.02$, $t(94) = 1.42$, $p = 0.16$.

In addition to valence, the characteristics of the information source can have an impact on how the information is perceived. Interestingly, the negativity bias literature seems to completely overlook the influence of information sources. In the two widely cited literature survey papers by

Baumeister et al. (2001) and Rozin and Royzman (2001), the authors critically review important works on negativity bias in past decades, but neither discuss the relationship between information source, information valence, and the perceived value of the information. On the other hand, source identity and source reputation have been acknowledged as key factors in health communications (e.g., Eastin, 2001), marketing (e.g., Homer & Kahle, 1990), and persuasive communications generally (e.g., DeBono & Harnish, 1988). Also, recent studies of online reviews have found that online review readers use source cues (such as reviewers' identity disclosures, and community-rated reputation) to help filter and process review content (Forman et al., 2008; Hu et al., 2008; Vermeulen & Seegers, 2009). These studies generally show that source identity disclosure and source reputation have significant effects on evaluations of information in various contexts. Therefore, the present study introduced the reputation of reviewers as a moderator, which has two conditions: high and low. Following the reputation mechanism on TripAdvisor, a "Top Contributor" icon was used to label the high-reputation reviewer, and a simple "Reviewer" title identified the low-reputation reviewer. A "Top contributor" is someone who has posted a large number ofrestaurant reviews in the past, and has received many helpfulness votes from others; "Reviewer" is someone who has posted a small number of reviews in the past and received no helpfulness votes. The reputation labels, number of past reviews, and number of helpfulness votes were displayed under each reviewer's screen name. An illustration of the artificial reviews along with the reviewer reputation details is included in Appendix 2.

To conduct the experiments for this research, the researcher registered as a requester on MTurk and posted the HIT in November 2011. In the HIT descriptions, the researcher described

the task as part of an academic research project about online reviews, but did not explicitly state

the central theme of the research (i.e., negativity bias). The workers who completed the HITs

were paid $0.30, which was comparable to other HIT payments on MTurk (e.g., Berinsky,

Huber, & Lenz, 2012; Downs, Holbrook, Sheng, & Cranor, 2010). Using MTurk's prescreening

feature, HITs were accessible only to those who had a high HIT approval rate (higher than 95%),

who were aged 18 years or older, and who lived in the United States. However, these preliminary

measures might not be sufficient to prevent participants from "gaming" the system, since cash

payout and anonymity might entice some workers to complete as many HITs as possible without

fully engaging in them. Following Downs et al. (2010), the experiment instruments included a

screening question to identify non-engaged workers who mighthave produced useless data.

Because MTurk provides only a rudimentary survey tool, the experimental instrument was

administered through a Web-based survey platform (Qualtrics). Upon agreeing to accept the

HITs on the MTurk website, workers were directed to a Qualtrics webpage. Using the

randomizer on Qualtrics, workers were randomly assigned to different conditions. Qualtrics was

able to track the respondents' IP addresses so that only one participation from each IP was

allowed for one HIT. Upon completing the Qualtrics survey, respondents were assigned a unique

code and instructed to go back to MTurk to claim their payment by supplying their MTurk user

ID and this code.

Participation was 292 MTurk workers in this 3 (valence: positive, mixed, negative) x 2

(reputation: high, low) between subjects experiment. Participants were randomly assigned to one

of the six conditions using the conditional branching function of the online survey software.

Each subject was presented with a webpage that resembled a screen shot of a typical online

review page that included the reviewer's screen name, the review text, and the number of stars (out of 5) summarizing the overall valence of the review. After reading the review page, the participant rated the helpfulness of the review on three items ("informative", "useful", "helpful"), using a 7-point Likert scale (1 = "not at all" and 7 = "very"). As already mentioned, a screening question was included in the questionnaire to identify non-engaging participants. The screening question asked which food item out of four answer choices was not mentioned in the restaurant review. As a manipulation check, participants were asked to rate the valence of each review using a 5-point scale. The questionnaire ended with a set of demographic items, and participants received a HIT completion code before exiting the survey software.

All 292 participants submitted the correct completion code when claiming payment on MTurk. However, 21 respondents failed to answer the screen question correctly, and 5 did not complete the helpfulness questions. These participants were excluded from the data analysis, which resulted in a usable dataset of 266 responses. A composite index of the overall helpfulness of the reviews was created by combining the three helpfulness items ($\alpha = 0.92$). The ANOVA of helpfulness demonstrated that the main effects of valence ($F_{(2, 260)} = 12.09$, $p < 0.001$) and reputation ($F_{(1, 260)} = 5.88$, $p < 0.05$) are both significant. A post-ANOVA Turkey HSD pair-wise comparison revealed that positive and negative reviews were both perceived as more helpful than mixed reviews ($p < 0.01$ and $p < 0.001$, respectively), while the difference between positive and negative reviews was not significant ($p = 0.36$). The results show no statistically significant interaction effect, either ($F_{(2, 260)} = 1.85$, $p = 0.16$).

[Insert Figure 1 about here]

Study 2 confirms the results from Study 1 in that negativity bias is not evident when

consumers evaluate the helpfulness of online reviews. Positive reviews were rated as helpful as

negative reviews, and more helpful than mixed reviews. In other words, there is no linear and

positive association between negativity and helpfulness of the reviews. The negative and the

mixed reviews from the high-reputation reviewer were perceived as slightly more helpful than

the reviews from the low-reputation reviewer. However, the ANOVA shows no significant

moderating effect of reviewer reputation.

Study 3

The results of Studies 1 and 2 are consistent but somewhat surprising given the commonly

held belief that "bad is stronger than good" (Baumeister et al., 2001). A possible explanation

might be that the so-called negativity bias is not an inherent perceptual bias but a context-

dependent phenomenon resulting from the rarity of negative events. The infrequency of negative

events makes negative things more attention-grabbing and more diagnostic (Taylor, 1991). In

general, in daily life, positive events occur more often than negative ones, resulting in a "positive

baseline" against which valence judgments are made. Indirect evidence comes from Boucher and

Osgood's (1969) study of 13 languages, which finds a universal human tendency in

communication to use evaluatively positive words more frequently than evaluatively negative

words. Data analysis of Study 1 reveals also that both the book reviews and the helpfulness

ratings of the reviews lean towards the positive. Since the majority of online reviews are

positive, negative reviews tend to carry more weight and be perceived as more helpful. Study 2

participants were exposed to equal numbers of positive and negative reviews of an unfamiliar restaurant, which was an artificial context lacking a positive baseline.

To investigate whether a baseline condition would attenuate or even reverse the negativity bias, the researcher used the same positive and negative restaurant reviews as in Study 2 but presented them in two different baseline contexts. In the positive baseline context, the fictitious restaurant XYZ Kitchen had an average customer rating of 4.5 (out of 5) stars and was "ranked # 9 of 8,600 restaurants in New York City"; in the negative baseline context, the restaurant received an average customer rating of 2 stars and was "ranked #6132 of 8,600 restaurants in New York City". Aggregate star ratings and overall ranking of a product in a category provide consumers with a quick overview of the overall evaluative sentiments of other consumers. These two features are commonly seen on mainstream e-commerce websites including TripAdvisor and Amazon. Each of the two valenced reviews used in Study 2 was presented in one of the two baseline conditions. This resulted in a 2 (baseline valence: positive, negative) x 2 (review valence: positive, negative) factorial design. Appendix 3 depicts the baseline conditions for the negative review.

Participation in this experiment was 205 MTurk workers. To eliminate potential carry-over effects, seven responses were excluded from the analysis because these respondents had participated in Study 2 (based on MTurk ID matching). A similar screening question and the demographics items used in Study 2 were included in this study. Six respondents who failed to answer the screening question correctly were excluded from the analysis. The remaining 192 participants were randomly assigned to one of the four conditions using the conditional

branching function of the Qualtrics survey tool. Each participant was instructed to rate the helpfulness of the review using the same measurement items in Study 2 ("informative", "useful", "helpful" on 7-point Likert scales).

A composite index of the overall helpfulness of the reviews was created by combining the three helpfulness items ($\alpha = 0.89$). A 2×2 ANOVA test was conducted with review valence and baseline condition as the independent variables, and review helpfulness as the dependent variable. There were no significant main effects ($F_{valence} = 1.23$, $F_{baseline} = 0.17$), but there was a significant interaction ($F(1,187) = 9.93$, $p < .01$). The negative review was rated higher for helpfulness under the positive baseline condition ($M = 5.60$) than under the negative baseline condition ($M = 5.02$; $t < 0.05$). Conversely, the positive review was rated more helpful under the negative baseline condition ($M = 5.72$) than under the positive baseline condition ($M = 5.27$; $t < 0.05$).

[Insert Figure 2 about here]

The results provide strong support for the hypothesized context-dependency of negativity bias. Participants only rated the negative review as more helpful than the positive review when the overall valence of consumer opinions was positive. When overall valence switched to negative, the so-called negativity effect was reversed and the positive review was rated more favorably than the negative review for helpfulness. The perceived helpfulness of online reviews in this case seems to derive from the novelty and surprisingness of the information rather than the individual review's valence.

DISCUSSION

In contrast to common wisdom and previous academic studies, this research found no empirical evidence of negativity bias when evaluating the helpfulness of online reviews. Study 1 showed that the valence of Amazon book reviews is positively associated with the helpfulness of the reviews. This correlation remains significant even after taking account in the regression model of review length and readability. A possible explanation is that satisfied customers are motivated to write well-composed and in-depth reviews, while unhappy customers use the reviews to vent their frustration, and provide less transferable information. Thus, , the study indicates that the valence of a customer review is less important than the quality of the information provided in the review. From the perspective of information diagnosticity, Study 1 confirms the critical importance of information quality for determining the usefulness to consumers of a piece of information. It is also logical to expect that the richer the information contained in a review, the more helpful it will be to other consumers. Negative information might grab the attention more easily, but attention alone does not guarantee the value of the information.

When the amount and the quality of information were controlled for in Study 2, the negative reviews and the positive reviews were rated equally helpful. This finding corroborates the findings from Study 1 in that both failed to reproduce the negativity effect documented in the psychology literature. Study 2 showed also that reviewer reputation has little influence on the perceived helpfulness of reviews since the association between review valence and perceived

helpfulness was not in any way moderated by reviewer reputation.

To try to explain the non-existence of the widely recognized perceptual negativity bias, Study 3 considered the baseline context of people's value evaluations. Past research shows that people's psychological anchors for value judgments tend to be at the positive end of the judgment scale and to be moderate in extremity (Skowronski & Carlston, 1989). That is, people are inclined to be "nice", and to give more positive evaluations when judging something or someone. Study 1 in this research, and other academic work on eWoM (e.g., Hu, Zhang, & Pavlou, 2009), reveal a binominal distribution of eWoM, with positive evaluations the most common. Precisely because people generally expect others to be moderately positive, negative information is likely to be perceived as novel and more valuable (Fiske, 1980). However, by manipulating the baseline valence conditions, Study 3 demonstrates that the perceived helpfulness of online reviews is more likely a result of a novelty effect rather than a negativity effect.

The findings from the three empirical studies show consistently that the negativity bias documented in the psychology literature may not be so applicable to the context of eWoM. In other words, bad is not necessarily stronger than good in relation to the perceived value of consumer-generated reviews. This conclusion encourages a new theorization of WoM through an exploration of the qualitative characteristics of WoM messages in addition to their valence. More specifically, this research highlights that the amount and quality of information contained in a review are critical for determining the perceived helpfulness of the review. An important point worth restating is that prior research in the marketing field tends to focus on the effect of

negative WoM on customers' brand evaluations and purchase intentions (e.g., C. Park & Lee, 2009) rather than the value of the WoM messages as perceived by consumers. It may be that negative WoM does have a stronger influence on sales than positive WoM, but the valence of a message becomes less relevant when it provides customers with useful and novel information.

Given the importance of eWoM's information quality, e-commerce companies need to consider mechanisms to encourage not only more positive but also more information-rich customer reviews that will be helpful to future customers. For example, websites such as Amazon could include readability assessment tools showing readability scores in real-time, while the customer is writing his or her review. In addition, information quality criteria could be used to order the customer reviews that appear on product pages so that potential customers could spot more useful reviews more quickly.

Marketing practitioners might want to pay special attention to the order and presentation of valenced customer reviews on websites. While multiple positive reviews are likely to create a favorable impression of a product, the informational value of the reviews is diminished by each review providing the same diagnostic cue for consumers' decision-making. A negative review in an overwhelmingly positive eWoM context would provide novel information, much needed by and helpful to consumers. More interestingly, negative details in a generally positive product evaluation might foster more positive product evaluations since the negative information creates a blemishing effect caused by the bad highlighting the salience of the good (Ein-Gar, Shiv, & Tormala, 2012). On the other hand, in an overall negative WoM context, a well-versed and detailed positive review could counterbalance the negativity effect and help the consumer to

make an informed decision. Future research should examine the influence of various mixes of positive and negative reviews with different qualities, on consumers' attitudes to products.

CONCLUSION

By combining data mining and virtual experimentation techniques, this research demonstrated that the widely held belief that "bad is stronger than good" requires critical scrutiny when applied to marketing practices. In a generally positive baseline context, a piece of negative information is likely to be perceived as novel, and therefore will capture people's attention more easily. However, attention alone does not guarantee the perceived value of the information. In addition, the novelty and unexpectedness of negative information can diminish when the overall valence in the evaluative context turns negative. These observations echo Ahluwalia's (2002) conclusion that the negativity effect in product evaluation is overstated in the consumer behavior literature.

While there is a growing body of literature on eWoM in marketing and other management disciplines, most researchers adopt a data mining approach and analyze textual data collected from the Internet. It is hoped that the current work will inspire more experimental studies that will lead to more solid theoretical accounts. Further research is needed to investigate various conditions and contexts in which a negative effect may or may not occur. It is important also that management researchers revisit the psychology literature on negativity effects in order to develop a comprehensive understanding of different types of negativity bias in current consumer environments. For example, while this research focused on informational negativity bias in terms

24

of the perceived value of a piece of information, future studies could examine the persuasiveness

of negative information in relation to influencing attitude formation and actual behavior.

Appendix 1

An Example of Customer Review on Amazon.com

12 of 17 people found the following review helpful:

⭐⭐⭐☆☆ **Lots of Name-Dropping - and Good Stories**, February 6, 2011

By **W. A. Carpenter** (Portland, OR USA) - See all my reviews
VINE™ VOICE   REAL NAME

This review is from: Tell to Win: Connect, Persuade, and Triumph with the Hidden Power of Story (Hardcover)

Customer review from the Amazon Vine™ Program (What's this?)

One suspects that Tell to Win by Peter Guber was dictated by the author. The stories, most of which are very entertaining, seem like well-practiced anecdotes that would be great delivered in person but which lose a little something in print.

And practically every story involves one or more (sometimes many more) famous and important people - there is a lot of name dropping in this book.

Each chapter ends with a short "aHHa!" section that summarized key points of story telling as illustrated in that chapter. It would be difficult for someone who is not skilled at story telling to get better by following these principles.

Ultimately that is the key problem with this book. The author persuades us that telling stories is an important way to persuade other people but he doesn't teach how to tell good stories.

Help other customers find the most helpful reviews   Report abuse | Permalink

Was this review helpful to you? (Yes) (No) | 💬 Comment (1)

Artificial Customer Reviews for Study 2

Low-reputation condition:

Excellent food & service!
★★★★★

Reviewer: imfoodie

☆ 3 reviews 🎗 0 helpful vote

Reviewed on November 28, 2011

My friend and I stopped in this restaurant by chance and had an amazing experience. It was busy lunch time, but we were lucky enough to get a table immediately. Our waiter was attentive and knowledgeable. The menu had a nice variety of selections and specials. I ordered sea bass and my friend had filet mignon. Everything was cooked to perfection and tasted delicious. It was perhaps the best sea bass that I ever had! All in all, highly recommended!

OK but nothing special
★★★☆☆

Reviewer: imfoodie

☆ 3 reviews 🎗 0 helpful vote

Reviewed on November 28, 2011

My friend and I stopped in this restaurant by chance. The food was good quality, but there was not much memorable or unique about the restaurant. It was lunch time and we had to wait a short while for a table. The menu had more or less ordinary dishes. Service was attentive but somewhat robotic. I ordered sea bass and my friend had filet mignon. The sea bass was a bit dry, and my friend said his beef was just OK. Overall, not a bad restaurant, but there are many better ones in town.

Terrible food & service!

★☆☆☆☆

Reviewer: imfoodie

☆ 3 reviews 🏅 0 helpful vote

Reviewed on November 28, 2011

My friend and I stopped in this restaurant by chance. The place looked nice from outside so we thought it might worth a try. HUGE mistake! The restaurant was crowded and we waited 45 minutes to get a table. The waiter rushed us over to a table and then neglected us for a long while. Limited menu with poor selections. I ordered sea bass and my friend had filet mignon. I didn't finish the sea bass as it was dry and tasteless. My friend had to send back his steak because it was overcooked. An awful dining experience and I'll never go back!

High-reputation condition:

Terrible food & service!

★☆☆☆☆

Reviewer: imfoodie   🏅 **Top Contributor**

☆119 reviews 🏅 112 helpful votes

Reviewed on November 28, 2011

My friend and I stopped in this restaurant by chance. The place looked nice from outside so we thought it might worth a try. HUGE mistake! The restaurant was crowded and we waited 45 minutes to get a table. The waiter rushed us over to a table and then neglected us for a long while. Limited menu with poor selections. I ordered sea bass and my friend had filet mignon. I didn't finish the sea bass as it was dry and tasteless. My friend had to send back his steak because it was overcooked. An awful dining experience and I'll never go back!

Baseline Conditions for the Negative Review in Study 3

Positive baseline condition:

XYZ Kitchen

Address: 550 E 58th St, New York City, NY 10022

Cuisines: American, Mediterranean, Central European
Dining options: Breakfast/Brunch, Lunch, Dinner, Reservations, After-hours
Dining style: Casual Dining

XYZ Kitchen's simple-yet-inspired American-Mediterranean cuisine, warm hospitality, and a friendly and professional service team, helps create the quintessential New York dining experience for New Yorkers and visitors alike. Advance reservations recommended during busy hours.

Average Customer Reviews ★★★★½
Ranked # 9 of 8,600 restaurants in New York City

The latest customer review:

Terrible food & service!

★☆☆☆☆

Reviewed on May 28, 2013
My friend and I stopped in this restaurant by chance. The place looked nice from outside so we thought it might worth a try. HUGE mistake! The restaurant was crowded and we waited 45 minutes to get a table. The waiter rushed us over to a table and then neglected us for a long while. Limited menu with poor selections. I ordered sea bass and my friend had filet mignon. I didn't finish the sea bass as it was dry and tasteless. My friend had to send back his steak because it was overcooked. An awful dining experience and I'll never go back!

Negative baseline condition:

XYZ Kitchen

Address: 550 E 58th St, New York City, NY 10022

Cuisines: American, Mediterranean, Central European
Dining options: Breakfast/Brunch, Lunch, Dinner, Reservations, After-hours
Dining style: Casual Dining

XYZ Kitchen's simple-yet-inspired American-Mediterranean cuisine, warm hospitality, and a friendly and professional service team, helps create the quintessential New York dining experience for New Yorkers and visitors alike. Advance reservations recommended during busy hours.

Average Customer Reviews ★★☆☆☆
Ranked # 6132 of 8,600 restaurants in New York City

The latest customer review:

Terrible food & service!
★☆☆☆☆
Reviewed on May 28, 2013
My friend and I stopped in this restaurant by chance. The place looked nice from outside so we thought it might worth a try. HUGE mistake! The restaurant was crowded and we waited 45 minutes to get a table. The waiter rushed us over to a table and then neglected us for a long while. Limited menu with poor selections. I ordered sea bass and my friend had filet mignon. I didn't finish the sea bass as it was dry and tasteless. My friend had to send back his steak because it was overcooked. An awful dining experience and I'll never go back!

References

Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of Consumer Research*, *29*(2), 270–79.

Amabile, T. M. (1983). Brilliant but cruel: Perceptions of negative evaluators. *Journal of Experimental Social Psychology*, *19*(2), 146–156.

Amabile, T. M., & Glazebrook, A. H. (1982). A negativity bias in interpersonal evaluation. *Journal of Experimental Social Psychology*, *18*(1), 1–22.

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of Marketing Research*, *4*(3), 291–295.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing*, *67*(4), 103–117.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

Berger, J., Sorensen, A. T., & Rasmussen, S. J. (2010). Positive effects of negative publicity. *Marketing Science*, *29*(5), 815–827.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351–368.

Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *8*(1), 1–8.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5.

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

Danescu-Niculescu-Mizil, C., Kossinets, G., Kleinberg, J., & Lee, L. (2009). How opinions are received by online communities: A case study on Amazon.com helpfulness votes. *Proceedings of the 18th International Conference on World Wide Web (WWW'09)* (pp. 141–150). Madrid, Spain: ACM Press.

DeBono, K. G., & Harnish, R. J. (1988). Source expertise, source attractiveness, and the processing of persuasive information: A functional approach. *Journal of Personality and Social Psychology*, *55*(4), 541–546.

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

Dellarocas, C., Zhang, X., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45.

Dowling, G. R., & Staelin, R. (1994). A model of perceived risk and intended risk-handling activity. *Journal of Consumer Research*, *21*(1), 119–134.

Downs, J. S., Holbrook, M. B., Sheng, S., & Cranor, L. F. (2010). Are your participants gaming the system? Screening Mechanical Turk workers. *Proceedings of the 28th international conference on Human factors in computing systems (CHI2010)* (pp. 2399–2402). New York, NY, USA: ACM.

Duan, W., Gu, B., & Whinston, A. B. (2008a). Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems*, *45*(4), 1007–1016.

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

Eastin, M. S. (2001). Credibility assessments of online health information: the effects of source expertise and knowledge of content. *Journal of Computer‐Mediated Communication*, *6*(4), 0–0.

Ein-Gar, D., Shiv, B., & Tormala, Z. L. (2012). When blemishing leads to blossoming: The positive effect of negative information. *Journal of Consumer Research*, *38*(5), 846 – 859.

Eisend, M. (2013). The moderating influence of involvement on two-sided advertising effects. *Psychology & Marketing*, *30*(7), 566–575.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*(6), 889–906.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221–233.

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do users evaluate the credibility of Web sites? A study with over 2,500 participants. *Proceedings of the 2003 conference on Designing for user experiences* (pp. 1–15).

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291–313.

Ghose, A., & Ipeirotis, P. (2009). The EconoMining project at NYU: Studying the economic value of user-generated content on the internet. *Journal of Revenue and Pricing Management*, *8*(2-3), 241–246.

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

Hennig-Thurau, T., & Walsh, G. (2004). Electronic word of mouth: Motives for and consequences of reading customer articulations on the Internet. *International Journal of Electronic Commerce*, *8*(1), 51–74.

Herr, P. M., Kardes, F. R., & Kim, J. (1991). Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective. *Journal of Consumer Research*, *17*(4), 454–462.

Homer, P. M., & Kahle, L. R. (1990). Source expertise, time of source identification, and involvement in persuasion: An elaborative processing perspective. *Journal of Advertising*, *19*(1), 30–39.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425.

Hu, N., Liu, L., & Zhang, J. J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management*, *9*(3), 201–214.

Hu, N., Zhang, J., & Pavlou, P. A. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM*, *52*(10), 144–147.

Huang, P., Lurie, N. H., & Mitra, S. (2009). Searching for experience on the web: An empirical examination of consumer behavior for search and experience goods. *Journal of Marketing*, *73*(2), 55–69.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*(4), 887–900.

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, *44*(2), 544–564.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel.* National Technical Information Service, Springfield, Virginia. Retrieved July 26, 2013, from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED108134

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

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, *58*(13), 2078–2091.

Mudambi, S., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200.

Naylor, R. W., Lamberton, C. P., & Norton, D. A. (2011). Seeing ourselves in others: Reviewer ambiguity, egocentric anchoring, and persuasion. *Journal of Marketing Research*, *48*(3), 617–631.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419.

Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business Research*, *62*(1), 61–67.

Park, D.-H., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, *11*(4), 125–148.

Park, S.-B., & Park, D.-H. (2013). The effect of low- versus high-variance in product reviews on product evaluation. *Psychology & Marketing*, *30*(7), 543–554.

Peeters, G. (1971). The positive-negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, *1*(4), 455–474.

Richins, M. L. (1983). Negative word-of-mouth by dissatisfied consumers: A pilot study. *The Journal of Marketing*, *47*(1), 68–78.

Riskey, D. R., & Birnbaum, M. H. (1974). Compensatory effects in moral judgment: Two rights don't make up for a wrong. *Journal of Experimental Psychology*, *103*(1), 171–173.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296 –320.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*(1), 131–142.

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 780–805.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*(1), 67–85.

De Valck, K., van Bruggen, G. H., & Wierenga, B. (2009). Virtual communities: A marketing perspective. *Decision Support Systems*, *47*(3), 185–203.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, *30*(1), 123–127.

Weathers, D., Sharma, S., & Wood, S. L. (2007). Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods. *Journal of Retailing*, *83*(4), 393–401.

Wu, P. F., van der Heijden, H., & Korfiatis, N. (2011). The influences of negativity and review quality on the helpfulness of online reviews. *Proceedings of the International Conference on Information Systems (ICIS2011)*. Atlanta: Association for Information Systems.

Table 1. Descriptive statistics

Variables	Minimum	Maximum	Mean	SD
Book rating	1	5	4.15	1.13
Total helpfulness votes	0	2103	8.75	48.44
Number of "helpful" votes	0	1934	5.73	41
Ratio of "helpful" votes	0	1	0.64	0.37
Length of review (in words)	1	5658	155.51	168.62
Review's FRE readability index	0	100	67.35	11.27

Note: N = 40,755

Table 2. OLS regression results

Model	Independent Variable	Coefficient	t	Sig.
1	Valence	0.288	8.317	**< 0.001**
	Valence^2	-0.188	-5.416	**< 0.001**
2	Valence	0.170	4.973	**< 0.01**
	Valence^2	-0.050	-1.452	0.146
	FRE	- 0.063	-11.110	**< 0.001**
	WordCount	0.157	27.623	**< 0.001**

Table 3. Summary of stepwise regression model

Model	R	R^2	Std. Error	ANOVA	
				F Change	Sig.
1	0.107	0.012	0.258	187.066	< 0.001
2	0.210	0.044	0.250	372.038	< 0.001

Figure 1. Review helpfulness as a function of reviewer reputation and review valence
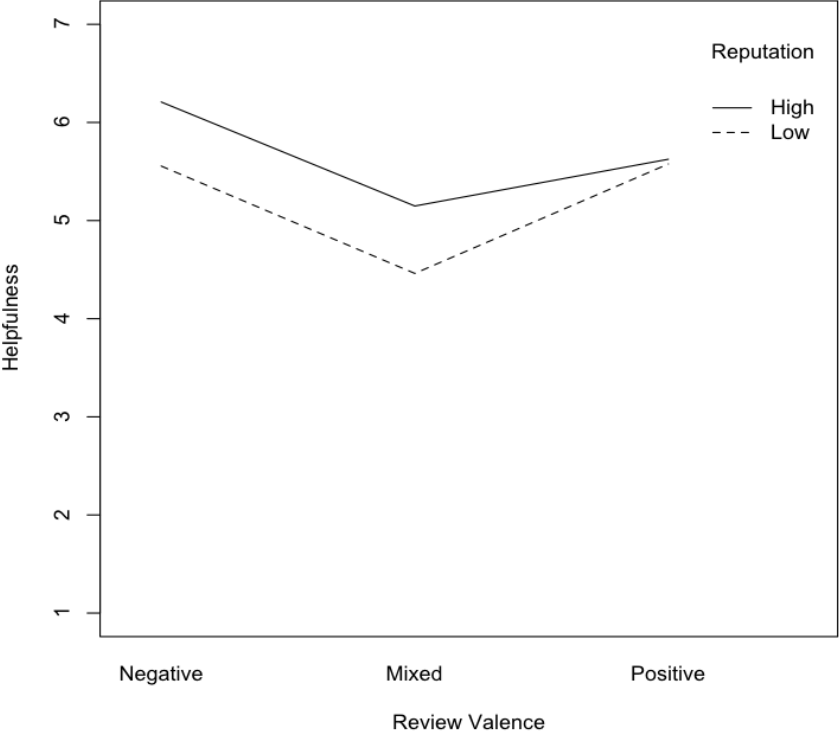
Figure 2. Review helpfulness as a function of baseline condition and review valence