

**Keep Your Cool or Let It Out:
Nonlinear Effects of Expressed Arousal on Perceptions of Consumer Reviews**

Dezhi Yin

Assistant Professor of Management
Trulaske College of Business, University of Missouri
440 Cornell Hall, Columbia, MO 65211, U.S.A.
(573) 884-6087
yind@missouri.edu

Samuel D. Bond

Associate Professor of Marketing
Scheller College of Business, Georgia Institute of Technology
800 West Peachtree, NW, Atlanta, GA 30308, U.S.A.
(404) 385-3309
sam.bond@scheller.gatech.edu

Han Zhang

Associate Professor of Information Technology Management
Scheller College of Business, Georgia Institute of Technology
800 West Peachtree, NW, Atlanta, GA 30308, U.S.A.
(404) 894-4373
han.zhang@scheller.gatech.edu

Abstract: This research explores how expressed emotional arousal in a consumer review affects reader perceptions of its helpfulness. Drawing from research on written communication and lay theories of emotion, the authors propose a pattern of diminishing returns, in which the marginal effect of arousal on perceived helpfulness is positive at low levels of arousal but diminishes at higher levels. Results of a field study using Apple's App Store, a follow-up survey, and two laboratory experiments provide consistent evidence for the predicted pattern. In addition, results suggest that the nonlinear effect is explained in part by perceptions of reviewer effort, and that the effect is stronger for products that are utilitarian in nature. By revealing a nuanced relationship between emotional expression and perceived helpfulness, these findings offer valuable implications for effective word-of-mouth communication.

Keywords: Word-of-mouth, Arousal, Emotion, Helpfulness, Consumer reviews

Acknowledgement: The authors thank Sabyasachi Mitra, Sandra Slaughter, Jack Feldman, Detmar Straub, Eric Overby, Xiao Huang, and Lin Jiang for their insightful feedback on earlier versions of the paper. The authors are grateful to Marius Florin Niculescu, Lizhen Xu, and Michael Cummins for their generous help in recruiting experiment participants. Finally, the authors thank the editor, associate editor, and three anonymous reviewers for valuable guidance throughout the review process.

***** Forthcoming at *Journal of Marketing Research* *****

The dramatic growth of Internet-enabled technologies has magnified the ability of consumers to influence one another through direct and indirect communication. Reflecting this trend, online review platforms allow consumers to exert considerable influence by sharing their opinions of products and services (Chen and Xie 2008; Chevalier and Mayzlin 2006; Dellarocas 2003). Given the sheer quantity of reviews available, many platforms offer mechanisms to identify or promote reviews considered ‘helpful.’ However, current understanding of factors influencing the perceived helpfulness of reviews is limited. Greater understanding of these factors would benefit review platforms and consumers themselves.

Past research on antecedents of review helpfulness has highlighted factors such as numerical ratings, review timing, and product or reviewer characteristics (Chen and Lurie 2013; Forman, Ghose, and Wiesenfeld 2008; Mudambi and Schuff 2010; Yin, Mitra, and Zhang 2016). More recently, an emerging stream has begun to explore verbal expression of emotion, which is highly prevalent in reviews and can exert substantial impact on reader judgments (Schindler and Bickart 2012). The vast majority of work in this domain has focused on the valence of reviewer-expressed emotions. For example, Chen and Lurie (2013) demonstrated that perceptions of review helpfulness reveal a negativity bias driven by reader attributions for review content. Ludwig and colleagues (2013) showed that the positive or negative tone of online retailer reviews influences ultimate conversion rates. In contrast, the present research focuses on expression of emotional arousal, defined as the level of energy characterizing an emotional experience (Niedenthal 2008; Russell 1980). We examine how the emotional arousal expressed in a review impacts reader perceptions regarding its helpfulness.

Online platforms offer varying advice regarding emotional expression. The review site Yelp discourages reviewers from exaggerating their feelings, classifying “rants and raves” as

unhelpful (Yelp 2013); in contrast, Amazon encourages reviewers to share all their “energy and enthusiasm (both favorable and critical)” about a product (Amazon 2016). This disparity reflects broader disagreement about the role of emotions in judgment and decision-making. Conventional wisdom often treats emotion as an impediment to rationality, and ample evidence shows that intense emotional experiences can interfere with deliberation, self-control, etc. (Johnson and Tversky 1983; Loewenstein 1996; Shiv and Fedorikhin 1999; Tice, Bratslavsky, and Baumeister 2001). However, emotions also serve an important informational and motivational role, so that their impact on cognitive tasks is variable and sometimes beneficial (Damasio 2005; Isen 2001; Lerner and Keltner 2000; Naqvi, Shiv, and Bechara 2006).

Although there are myriad ways in which emotional arousal in a review may affect reader perceptions, we focus on a pathway of particular interest – inferences regarding reviewer *effort*. When utilizing prior reviews to inform their own decisions, consumers exhibit a tendency to “fill in blanks” by making inferences about reviewers (Naylor, Lamberton, and Norton 2011). An especially relevant inference concerns the amount of *effort* exerted by reviewers in deliberating on their experience and constructing a thoughtful review (Yin, Bond, and Zhang 2014). Building on contemporary models of emotion communication (Harris and Paradise 2007; Planalp 1998), we argue that readers make inferences of reviewer effort based on emotional arousal, which represents a fundamental dimension of emotional expression. Specifically, we theorize a nonlinear, diminishing-returns relationship between expressed emotional arousal and perceived review helpfulness, such that at low levels of expressed arousal, readers associate additional arousal with greater effort and a more helpful review, but at higher levels the effect of additional arousal diminishes, and may even become negative. Importantly, our prediction deals only with *perceptions* of reviewer effort and helpfulness (regardless of the accuracy of those perceptions).

To examine our prediction, we present a field study, a survey, and two laboratory experiments. Across our studies, expressed arousal is either measured or manipulated, in the form of verbal cues, non-verbal cues, or their combination. In addition, we explore perceived reviewer effort as a potential explanatory mechanism, and we examine the nature of the product (utilitarian or hedonic) as a theoretically relevant moderator (Moore 2015). By exploring the role of expressed arousal in consumer reviews, our research adds to a growing body of knowledge on emotions in WOM. Recent work has focused on the motivation of message recipients to share a message, based on the extent to which they find its content arousing (Berger 2011; Berger and Milkman 2012). We complement this work by focusing on inferences made by recipients about the sender and the message, based on the extent to which message content expresses arousal. More broadly, our exploration contributes to an important emerging literature on the social function of emotions (Van Kleef, De Dreu, and Manstead 2010).

THEORETICAL DEVELOPMENT AND HYPOTHESES

The Expression of Emotional Arousal

Psychologists have proposed a vast range of models to classify fundamental dimensions of emotional experience (Brosch, Pourtois, and Sander 2010; Mano 1991; Osgood 1966; Watson and Tellegen 1985). Although these models vary considerably, they consistently identify dimensions of valence and arousal to be particularly important (Niedenthal 2008; Russell 1980). Valence describes the extent to which an individual perceives an experience as pleasant or unpleasant and is used to distinguish ‘positive’ and ‘negative’ affective states. Arousal (also ‘activation’ or ‘intensity’) describes the extent to which an individual is energized by an experience. Scales measuring arousal utilize endpoints such as ‘calming’ vs. ‘exciting,’

‘soothing’ vs. ‘agitating,’ etc. (Bachorowski and Braaten 1994; Heilman 1997). Although valence and arousal sometimes co-vary, the two dimensions are independent and distinguishable, both phenomenologically and physiologically (Niedenthal 2008; Russell and Barrett 1999; Smith and Ellsworth 1985).

It is widely accepted that emotional response manifests through an “emotional reaction triad” of experience, physiology, and expression (Ekman 1993; Lang 1995; Levenson 1994; Scherer 2000). Applied here, experienced arousal captures the intensity of one’s subjective feelings, while physiological arousal captures bodily changes and physical responses: muscular tension, increased heart rate and blood pressure, etc. (Cannon 1927; Schachter and Singer 1962). In a consumer review setting, the experienced and physiological arousal of a review writer are not directly observable but must be inferred from the expression of arousal in the review content itself. Therefore, we focus on expressed arousal, and we use the terms “arousal” and “expressed arousal” interchangeably.

Emotional expression during communication occurs through verbal and nonverbal cues (Scherer 2000). Verbal cues represent the specific words embedded in a message; e.g., communicators may reveal emotions with descriptive words (“love”, “jealousy,” etc.) and convey emotional arousal with linguistic markers (“a bit”, “very”, “really,” etc.) (Harris and Paradice 2007). In contrast, nonverbal cues represent auxiliary signals that depend on the medium. In spoken communication settings, nonverbal cues may include facial expressions, vocal expressions, and body language. In written communication, a wide variety of nonverbal cues are available (Carey 1980; Planalp 1998): for example, writers may use emoticons (e.g., “^_^”) and other symbolic text to imitate facial or vocal expressions, or they may convey arousal through the use of grammatical markers. Capitalized words and exclamation marks that connote

high volume (“shouting”) are commonly used to indicate high-arousal emotions (Allen 1988; Schandorf 2013).

Arousal Cues and Reader Inferences

The online review setting represents an ‘actor-observer’ context, in which reviewers provide product-relevant content, and readers later make sense of both the content and the reviewer. Given widespread evidence that observers attend to and elaborate upon emotional cues (Van Kleef 2010), it is logical to expect that review-embedded emotions will be important in the sense-making process. Emotion-related words are often processed automatically, and perceivers tend to utilize emotional cues even when other sources of information are sufficient (Gendron et al. 2012; Gernsbacher, Hallada, and Robertson 1998). In computer-mediated communication (Harris and Paradice 2007; Vandergriff 2013), readers incorporate both verbal and non-verbal cues to assess the emotional state of senders, such that the presence of more such cues leads to assessments of greater sender arousal. Applied here, review readers who encounter more arousal cues are likely to infer that the reviewer experienced greater arousal during the process of constructing the review, and in turn generate related inferences about the reviewer and review.

Of numerous inferences relevant to our setting, we focus on inferences regarding reviewer effort. We define perceived reviewer effort as the extent to which a reader believes that a reviewer exerted thoughtful deliberation on the content and construction of his or her review (note that our definition does not involve effort expended in purchasing or using the product). Our focus on effort derives from both practical and theoretical considerations. Consumer processing of decision-relevant information typically involves an evaluation of the information source, and this is especially true in the case of product recommendations (Gershoff, Broniarczyk, and West 2001). Evidence from review contexts suggests that consumers form

spontaneous inferences about reviewer characteristics, which are typically unknown or ambiguous (Naylor, Lamberton, and Norton 2011). Moreover, consumers appear to recognize that the effort underlying a review is relevant to its usefulness, as objective indicators of reviewer effort predict helpfulness perceptions even after other factors are controlled for (Mudambi and Schuff 2010). Prior research indicates that consumers infer reviewer effort from the discrete emotions in a review (Yin, Bond, and Zhang 2014). Given that arousal is a fundamental component of emotional expression, it is reasonable to expect that perceived arousal may itself influence inferences regarding effort.

Research on the mental representation of emotion has identified commonly held beliefs regarding the characteristics, antecedents, and consequences of emotional experience (e.g., Shaver et al. 1987). Among these is the belief that emotional arousal serves an important motivational function. Individuals exhibiting low levels of arousal are commonly described by others as ‘relaxed,’ ‘calm,’ or ‘bored,’ indicating limited task-directed attention, information processing, or elaboration (O’hanlon 1981). In contrast, individuals exhibiting high levels of arousal are described by terms such as ‘excited’ or ‘enthusiastic,’ indicating readiness to expend energy and take action (Lang 1995; Seo, Barrett, and Bartunek 2004). This connection is corroborated by research on lay theories of specific emotions. For example, Frijda, Kuipers, and Ter Schure (1989) have shown that targets experiencing low-arousal emotions such as sadness or boredom are expected to generate little motivated action (‘apathy,’ ‘rest,’ etc.), while targets experiencing high-arousal emotions such as anxiety or anger are expected to engage in effortful exertion (e.g., ‘attending,’ ‘reactant,’ etc.). Applying this logic, we predict that up to a point, greater expressed arousal in a review will be associated by readers with greater effort expended in constructing the review.

Importantly, however, we also predict that the positive association between expressed arousal and perceived reviewer effort will weaken as arousal increases. At higher and higher levels of arousal, the level of implied effort will become more implausible, and readers will be increasingly likely to form alternative inferences. For example, readers encountering high levels of arousal may perceive that a reviewer was utilizing a more subjective vs. objective writing style, was simply ‘gushing’ or ‘venting’ about his or her experience (Jensen et al. 2013; Park, Lee, and Han 2007), or was engaging in strategic puffery – “trying too hard” to convince prospective consumers one way or the other (Friestad and Wright 1994; Xu and Wyer Jr 2010). In addition, readers may interpret extreme levels of arousal as a signal of irrationality. Although emotion is often an important input to reasoning tasks (Damasio 2005; Pham 2007), highly elevated arousal is associated with numerous cognitive shortcomings (Fedorikhin and Patrick 2010; Gorn, Pham, and Sin 2001; Pham 2007; Sanbonmatsu and Kardes 1988), and research on lay theories of emotion reveals a common belief that intense emotion is an impediment to reasoning (Oatley and Johnson-Laird 2014). At very high levels of expressed arousal, therefore, the relationship between arousal and perceived effort will weaken or even reverse.

Our final argument is that perceptions of reviewer effort are positively associated with perceptions of review helpfulness. This argument is both intuitive and consistent with broad interdisciplinary evidence that perceivers assume effort and performance to be correlated (see Skinner, Chapman, and Baltes 1988; Weiner and Kukla 1970). As a form of discursive writing (Vygotsky 1964), reviewing is deliberative in nature and involves various subtasks that each require cognitive resources: reconstructing one’s experience with a product, translating those thoughts into coherent form, incorporating assumptions about the audience, etc. Readers will naturally assume that reviewers who have expended greater effort performed these subtasks more

thoroughly (e.g., by reconstructing a more complete account of the experience), thereby producing more helpful reviews. Taken together, our arguments lead to the following:

Hypothesis 1: The level of expressed arousal in a review is associated with perceived review helpfulness in a nonlinear (diminishing returns) manner, such that the effect of additional arousal is positive at low levels of arousal, but becomes smaller as arousal increases.

Hypothesis 2: The nonlinear effect of expressed arousal on perceived review helpfulness is mediated by perceived effort, such that a) expressed arousal is associated with perceived reviewer effort in a nonlinear (diminishing returns) manner, and b) perceived reviewer effort is positively associated with perceived review helpfulness.

Arousal-Based Inferences across Product Categories

As typically defined, utilitarian products are primarily instrumental and consumed for functional purposes, while hedonic products are primarily experiential and consumed for pleasure (Dhar and Wertenbroch 2000; Sloot, Verhoef, and Franses 2005). Hedonic products tend to be more “affectively rich” than utilitarian products, and to evoke stronger emotional reactions (Babin, Darden, and Griffin 1994; Botti and McGill 2011). Given these intrinsic differences, it stands to reason that consumers will expect WOM for hedonic products to contain greater emotion than WOM for utilitarian products. Empirical evidence confirms that these expectations are often born out, as emotional language is more prevalent for online reviews in hedonic categories (Kronrod and Danziger 2013).

The tendency for reviews of hedonic products to be more emotional has important implications for reader inference. Research on affective decision making in intrapersonal contexts shows that emotions which “stand out” are more likely to be utilized in subsequent

decisions (Albarracín and Kumkale 2003; Greifeneder, Bless, and Pham 2011; Siemer and Reisenzein 1998). Applying similar logic to our setting, readers should perceive arousal cues to be more diagnostic for reviews of utilitarian products (where they are less common) than reviews of hedonic products (where they are more common). As a result, readers will be more likely to draw on those arousal cues to make inferences about the reviewer and review. Therefore, the impact of expressed arousal on perceived review helpfulness will be stronger for utilitarian products than hedonic products, resulting in a ‘steeper’ curve:

Hypothesis 3: The nonlinear effect of expressed arousal on perceived review helpfulness is greater for utilitarian products than hedonic products.

OVERVIEW OF STUDIES

We investigated our hypotheses with four studies using distinct methodologies. In Study 1, we gathered real-world evidence in the form of reviews at Apple’s ‘App Store.’ Helpfulness was measured directly by user votes, and expressed arousal was measured through textual analysis. In Study 2, we administered a survey in which respondents assessed 400 App Store reviews in terms of helpfulness, arousal, and reviewer effort. In the final two studies, we conducted laboratory experiments that permitted the direct manipulation of expressed arousal. Studies 1-3 also investigated the role of our proposed moderator, product type.

STUDY 1

Our first study utilized secondary data to test H1 and H3 in a real-world, online WOM setting. Data consisted of user reviews from Apple’s ‘App Store,’ which had accumulated nearly two years of user reviews at the time of data collection (April 2010). Visitors evaluate an app by

leaving a 1-to-5 “star” rating and writing a text review (see Figure 1). Below each review, the platform asks readers the question “Was this review helpful?” and provides ‘Yes’ and ‘No’ options. For reviews that have received at least one vote, both the number of ‘Yes’ votes and the number of total votes are displayed. The App Store platform provides developers a valuable channel for informing and persuading prospective customers. Moreover, the popularity of the platform ensures a wide range of apps and reviews, making it well suited to our investigation.

- Insert Figure 1 Here -

Data Collection

At the time of data collection, the App Store contained 20 mutually exclusive categories (“entertainment,” “productivity,” etc.). We began by identifying for each category apps ranked in the top 500 by popularity over the preceding three months. Of 62,266 apps identified in this manner, 40,417 had accumulated at least one review, and we retrieved all historical reviews for those apps. For each review, we collected the following information: rating, text review content, helpful votes (i.e., the number of readers who voted ‘Yes’), and total votes. We also collected app-level information including the following: average rating, count of all ratings, category, and whether or not the app was free. Prior to the analyses, we filtered reviews with non-English characters (94,815 total), reviews without any text (2,743 total), and reviews with a rating of ‘0’ (presumably due to system errors, 38 total). Of 1,623,497 reviews that remained, 418,415 had received at least one helpful vote, and these formed the pool for our analysis.

Variables

To measure the perceived helpfulness of a review, we used its total number of helpful votes as the dependent variable, with the total number of votes included as a covariate.¹

Importantly, the default ordering of App Store reviews during the period of data collection was by date posted rather than helpfulness, with more recent reviews appearing first.

Emotional valence and arousal were measured using text analysis software, the Revised Dictionary of Affect in Language (RDAL - Whissell 2009). The program has been widely used to quantify emotional content in psychology and linguistics, and substantial evidence supports its reliability and validity (Whissell 2009; Whissell et al. 1986). The RDAL dictionary contains 8,742 words characteristic of natural language and has been shown to match 90% of words in most samples. The dictionary includes not only emotional words but also words that appear commonly in natural language. Each word is assigned a score for both valence (1 = unpleasant, 3 = pleasant) and arousal (1 = passive, 3 = active). RDAL valence and arousal scores were developed by asking volunteers to rate random samples of words along one or the other dimension. For example, the arousal score for the word ‘run’ is substantially higher than that for ‘walk’ ($M = 2.8$ vs. 2.1), while valence scores for the two words are similar ($M = 2.2$ vs. 2.0).

To analyze the content of a text sample, RDAL matches each word in the sample with words in its internal dictionary; whenever a match is found, the scores for valence and arousal are retrieved. After all words in the sample have been analyzed, final scores for valence and arousal are computed by averaging retrieved scores for all identified words. Among the reviews in our set, more than 99% were assigned values for valence and arousal; words lacking either value were dropped from the analysis. Table 1 presents descriptive statistics, and Table 2 summarizes the operationalization of all variables.

- Insert Tables 1 and 2 Here -

To measure product type (utilitarian or hedonic) for each review, we conducted a pretest with 43 undergraduates. Participants were asked to evaluate each of the 20 App Store categories,

in random order. Instructions defined utilitarian apps as “useful, practical, functional, something that helps you achieve a goal or solve a specific problem” and experiential apps as “pleasant and fun, something that is enjoyable and appeals to your senses (the term "experiential" was used to aid comprehension). Each category name was presented along with a brief description of the category and names of several representative apps (see Web Appendix A). Participants answered the following, adapted from Sloot, Verhoef, and Franses (2005): “Using the scales below, how would you describe <category name> apps?” The question included two separate, 9-point items: “not utilitarian / very utilitarian,” and “not experiential / very experiential”. After compiling the results, we created a “utilitarian value” for each category by subtracting its average experiential rating from its average utilitarian rating.

As control variables in the analyses, we included three items that have been used in past literature on antecedents of review helpfulness (Korfiatis, Rodriguez, and Sicilia 2008; Mudambi and Schuff 2010): review rating, length,² and reading difficulty. Review rating indicates the star rating provided in a review; ratings ranged from 1 star to 5 stars ($M = 3.45$ stars). Review length indicates the number of words in a review ($M = 42$ words). Reading difficulty was calculated using the Gunning Fox Index (GFI, see Gunning 1969), which estimates the years of formal education needed to understand a piece of text on a first reading ($M = 7$ years). In addition, we controlled for the following characteristics at the app-level: average rating, count of ratings, and price. Average rating serves as an indicator of overall satisfaction, while the count of ratings indicates popularity. Price was captured with a dummy variable (0 = free, 1 = paid).

Data Analysis and Results

As a preliminary visual analysis, we created the chart in Figure 2, which depicts the relationship between review helpfulness and emotional arousal in the raw data. Consistent with predictions, the figure suggests a nonlinear trend of diminishing returns.

Given that the number of helpful votes was a count variable whose variance (50.47) exceeded its mean (2.37), negative binomial regression was utilized in the formal analysis. Results are presented in Table 3. To examine H1, we conducted a series of regressions in which we first entered control variables along with utilitarian value, expressed arousal, and squared arousal. As shown in Model 1, the coefficient of the squared term of arousal was negative and significant ($\beta = -.883, p < .01$), indicating a nonlinear relationship in the form of diminishing returns. Consistent with H1, increases in expressed arousal were associated with greater review helpfulness at low levels of arousal, but this positive association became weaker as arousal increased, and eventually became negative. Based on the obtained coefficients, the estimated inflection point was 1.69 on the 3-point arousal scale.

- Insert Figure 2 and Table 3 Here -

To investigate the role of product type, we created a new model (Model 2 in Table 3) which included the interaction of utilitarian value with both arousal and squared arousal. Consistent with H3, the estimated coefficient for the nonlinear interaction term was significant and negative ($\beta = -.027, p < .05$), indicating that the nonlinear effect of expressed arousal on perceptions of review helpfulness was greater for app categories that were relatively more utilitarian. The interaction is depicted graphically in Figure 3, which plots the expected value of review helpfulness against expressed arousal separately for more utilitarian categories (+1 SD) and more hedonic categories (-1 SD). The “steeper” curve for utilitarian vs. hedonic categories in the figure indicates support for our model.

- Insert Figure 3 Here -

Discussion

Using a sample of real-world reviews, Study 1 provided initial evidence that the relationship between the expressed arousal in a review and perceptions of its helpfulness follows a nonlinear pattern of diminishing returns, even after controlling for other variables. The impact of additional arousal was beneficial at lower levels of arousal, but this positive effect was reduced (and turned negative) at higher levels of arousal. Consistent with our inference-based account, the nonlinear effect of arousal was greater for utilitarian apps than hedonic apps.

However, the use of field data necessitated noteworthy limitations. First, expressed arousal was measured by software rather than human perceivers. Although useful, the RDAL is inherently constrained by its dictionary-matching algorithm, which ignores figurative language, grammatical markers, and other nonverbal characteristics (emoticons, etc.). Second, our framework proposes inferences regarding reviewer effort as a mechanism underlying the nonlinear impact of expressed arousal, but the design precluded measurement of such inferences. Study 2 was conducted to address these limitations. More broadly, the findings were correlational and thus subject to potential confounds; we address this concern with experiments in Studies 3-4.

STUDY 2

Our second study consisted of a survey in which respondents were given a set of real-world reviews, drawn from the pool in Study 1, and asked to provide their perceptions of each review. The design allowed us to investigate our primary hypothesis using a subjective measure

of arousal, and it also permitted examination of our proposed mediator, perceived reviewer effort.

Stimulus Materials and Procedure

We began with the full sample of reviews collected in Study 1. After dropping reviews less than one sentence long (to ensure enough content to generate an impression), we randomly selected 400 reviews to form the pool used in the study.

One hundred twenty-eight undergraduates participated in exchange for course credit. Participants were each shown 20 randomly selected reviews, one at a time and on separate screens. Each screen displayed the name of an app, its category, and the content of one review for that app. After reading the review, participants responded to the following measures: perceived helpfulness (two semantic differential items adapted from Sen and Lerman 2007; e.g., "not at all helpful / very helpful"); perceived arousal (two semantic differential items adapted from Berger 2011; e.g., "very mellow / very fired up"), perceived valence (two semantic differential items adapted from Berger 2011; e.g., "very negative / very positive"), and perceived effort (two Likert-type items adapted from Huddy, Feldman, and Cassese 2007; e.g., "In your opinion, how much thought did the reviewer give to this review? very little / very much"). Web Appendix C contains survey instructions and measures.

Data Analysis and Results

Because each participant evaluated different random reviews from the pool, analyses were conducted at the review level ($N = 400$). On average, each review was evaluated by 6.4 participants. For each review, we calculated average ratings for helpfulness, valence, arousal, and effort. Cronbach's alphas exceeded .90 for all constructs, indicating satisfactory reliability. Table 4 summarizes descriptive statistics and correlations of all major variables.

- Insert Table 4 Here -

To examine H1, we performed OLS regression by first entering arousal and all control variables, then adding a term for squared arousal. Results are depicted in Table 5 (Model 1). The coefficient for the squared term of arousal was negative and significant ($\beta = -.120, p < .01$). Consistent with H1, increased arousal was associated with greater review helpfulness at low levels of arousal, but the marginal effect of arousal became smaller as arousal increased, and eventually turned negative. The estimated inflection point was 5.89 on the 9-point arousal scale.

To investigate the role of perceived effort, we utilized the “MEDCURVE” macro (Hayes and Preacher 2010), which applies a non-parametric test that accounts for nonlinear relationships (Fritz and Mackinnon 2007). As illustrated in Figure 4, our nonlinear mediation model included a quadratic effect of arousal on perceived effort and a linear effect of perceived effort on review helpfulness, in addition to a quadratic direct effect of arousal on review helpfulness.

- Insert Table 5 and Figure 4 Here -

Analyses were conducted using 5,000 bootstrap resamples, and the control variables in Table 2 were included as covariates. Estimated model coefficients are presented in Figure 4. Supporting H2, results indicated the presence of nonlinear mediation. The quadratic effect of arousal on perceived effort was negative and significant ($\beta = -.108, p < .01$), and perceived effort was positively associated with review helpfulness ($\beta = .934, p < .01$). After controlling for the quadratic effect of arousal on perceived effort, the quadratic effect of arousal on helpfulness became non-significant ($\beta = -.019, p = .35$). To better understand the pattern of indirect effects, we calculated θ , the instantaneous indirect effect of expressed arousal on helpfulness through perceived effort, at low (-1 SD) and high ($+1$ SD) levels of arousal. Results revealed a significant positive indirect effect at low arousal ($\theta = .331, 95\% \text{ CI} = [.208, .491]$), suggesting

that perceived effort could explain the positive impact of arousal at low levels. Results also revealed a significant negative indirect effect at high arousal ($\theta = -.195$, 95% CI = $[-.381, -.005]$), suggesting that perceived effort could explain the negative impact of arousal at high levels.

To investigate the moderating role of product type, we constructed an additional model that included the interaction of utilitarian value obtained from Study 1 with both arousal and squared arousal (see Table 4, Model 2). The estimated coefficient for the nonlinear interaction term was not significant ($p > .4$). Thus, H3 was not supported.

Discussion

Using a survey design that allowed for subjective measures of our key variables, Study 2 replicated the main result of our field study: the relationship between expressed arousal and perceptions of review helpfulness followed a nonlinear pattern of diminishing returns. Consistent with our theorizing and H2, additional findings provided direct evidence that the nonlinear relationship was explained by perceptions of reviewer effort.

In contrast, Study 2 did not reveal evidence for the moderating role of product type. We speculate that this (non)result may be attributable to the review stimuli themselves, which were kept simple to reduce fatigue across 20 trials. Given that participants saw only the name of each app and its category, with no further description, they may not have been able (or motivated) to judge with confidence whether the app was utilitarian or hedonic in nature.

Neither of our first two studies provides definitive evidence of causation, due to the potential for unobserved confounds in the real-world reviews that they utilized. In particular, the level of arousal expressed in the reviews may have been associated with their information content, argument strength, etc.; if so, then the observed effects of arousal may simply reflect

differences in *objective* quality. We addressed these concerns in our next two studies by conducting laboratory experiments in which expressed arousal was manipulated directly.

STUDY 3

Study 3 examined expressed arousal and product type as manipulated (rather than measured) variables in a laboratory setting. Participants took part in a hypothetical shopping task in which they observed reviews for six different apps. The study incorporated a mixed design, with expressed arousal manipulated within-subjects at three levels (low, moderate, and high) and product type manipulated between-subjects at two levels (utilitarian, hedonic).

Stimulus Materials

We began by identifying real-world apps that were relatively unknown, relevant to a broad audience, and likely to evoke reviews varying in expressed arousal. To represent utilitarian and hedonic products, respectively, we narrowed our focus to the App Store ‘reference’ and ‘entertainment’ categories. From the reference category, we selected a location-based app that presents users with Wikipedia articles relevant to their surroundings. From the entertainment category, we selected a fish pond simulation, in which different varieties of fish swim on the screen and respond to user input.

The treatment reviews are presented in Table 6 and were developed in two stages. In the first stage, we created three ‘baseline reviews’ for each of the two apps. The baseline reviews were designed to provide one overall evaluative statement and one descriptive statement, with minimal emotional content. To construct the baseline reviews, we retrieved actual reviews from the App Store and then removed emotional words and phrases. For example, a baseline review

for the reference app was: “High quality app. It tells you how far away you are from the landmarks and even shows those landmarks on a map.”

- Insert Table 6 Here -

In the second stage, we varied each baseline review to create three treatment versions exhibiting low, moderate, or high arousal, through a combination of verbal cues and grammatical markers (Allen 1988; Schandorf 2013). The process involved four steps. First, we varied the presence and number of exclamation marks, such that that sentences in the low, moderate, and high arousal conditions ended with a period, an exclamation mark, or three exclamation marks, respectively. Next, we varied specific words in the text to convey different degrees of arousal (e.g., “noticed” vs. “was surprised” vs. “was completely shocked”). Third, we appended an emotional sentence to the moderate- and high-arousal reviews (e.g., “Happy with it!” vs. “Thrilled with it!!!”). Finally, we capitalized all letters in one word of each high-arousal review. The end result for each category was three sets of three reviews, varying in expressed arousal.

Procedure

Eighty-one undergraduate students participated in the study in exchange for course credit. Participants were randomly assigned to the utilitarian or hedonic category. In the cover story, they were asked to imagine that they were considering purchasing a specific variety of app (either “location-based” or “fish pond”), which was then briefly described (“learn about surprising new places and things around you”; “enjoy a beautifully rendered fish pond in your pocket”). Participants were informed that an initial search had returned six alternatives, with similar prices and similar average ratings (approximately four out of five stars). They were then told that to aid their decisions, they would be reading and evaluating reviews of each app.

The following screens presented six text reviews, one at a time. Participants were told that each review described a different app in their consideration set. Three ‘filler’ reviews were presented in positions 1, 3, and 5 in the sequence. The three treatment reviews were presented in positions 2, 4, and 6, and consisted of one low arousal, one moderate arousal, and one high arousal review. Treatment reviews were selected randomly from the three sets in Table 6, subject to the constraint that each review came from a different set, and they were presented in random order. This approach ensured that for each set, each level of arousal would be presented a similar number of times across participants.

After reading each review, participants were asked to report their perceptions of review helpfulness and perceived effort. As a manipulation check, participants also rated the level of arousal expressed in the review. Each variable was measured with three items, which included the two items used in Study 2 and one additional item (see Web Appendix D).

Results

Initial examination revealed satisfactory internal reliability for all measures (Cronbach’s alphas $> .9$). Analysis of the manipulation check revealed that perceived arousal in the high, medium, and low arousal conditions followed the intended pattern: $M = 8.26$ vs. 6.90 vs. 4.14 ; all $ps < .001$. Thus the manipulation of arousal was deemed successful.

To examine the effect of arousal on perceived review helpfulness, we conducted a mixed ANCOVA, with arousal entered as a within-subject factor, product type entered as a between-subject factor, and review order entered as a categorical covariate with six levels. The quadratic effect of arousal was specified as a polynomial contrast. Results of the analysis revealed no evidence for an overall linear effect of arousal ($p > .2$). Consistent with H1 and our first two studies, however, results revealed a significant quadratic effect of arousal ($F(1, 74) = 6.133, p <$

.05). As illustrated in Figure 5, pairwise comparisons revealed a significant increase in perceived helpfulness from low to moderate arousal ($M = 5.01$ vs. 5.63 , $t(80) = 2.63$, $p < .05$) and a non-significant decrease in perceived helpfulness from moderate to high arousal ($M = 5.34$, $t(80) = -1.29$, $p > .2$).

- Insert Figure 5 Here -

We next investigated the role of effort in mediating effects of arousal (H2). Because the study involved a repeated-measure variable with nonlinear effects, standard methods for assessing mediation like those of Study 2 were not applicable. Therefore, we adapted the three-step procedure proposed by Judd, Kenny, and McClelland (2001) for testing linear mediation in repeated-measure designs. Given that helpfulness did not significantly differ between moderate and high arousal, we restricted the analysis to low and moderate levels. First, repeated-measures ANCOVA revealed a positive effect of arousal on perceived effort ($M = 3.24$ vs. 4.26 , $t(80) = 5.41$, $p < .001$). Second, regression revealed that greater perceived effort was associated with greater perceived helpfulness at both low and moderate arousal levels ($\beta_s = .62$ and $.67$, $t_s = 5.63$ and 6.76 , $p_s < 0.01$). Third, differences in perceived effort predicted differences in perceived helpfulness ($\beta = .70$, $t = 6.79$, $p < .01$). Thus all three conditions were established, suggesting that the positive effect of arousal below moderate levels was mediated by perceived effort.

Finally, we investigated the potential moderating effect of product type. The interaction between the quadratic effect of arousal and product type did not reach conventional significance ($F(1, 74) = 2.680$, $p = .11$). For exploratory purposes, however, we conducted follow-up analyses for each product type separately. Results revealed a pattern consistent with Study 1: the quadratic effect of arousal was significant for the utilitarian category ($F(1, 34) = 7.249$, $p < .05$) but was not significant for the hedonic category ($F(1, 35) = .086$, $p > .7$). Within the utilitarian category,

pairwise comparisons revealed a significant increase in helpfulness perceptions from low to moderate arousal ($M = 5.39$ vs. 6.25 , $t(39) = 2.48$, $p < .05$), and a marginally significant decrease in helpfulness perceptions from moderate to high arousal ($M = 5.70$, $t(39) = -1.76$, $p = .09$).

Within the hedonic category, no pairwise comparison was significant ($ps > .4$).

Discussion

Extending our investigation to an experimental setting, Study 3 provided additional evidence that perceptions of review helpfulness are affected by expressed arousal in a manner of diminishing returns. In addition, mediation analyses supported our argument that the effect is due in part to inferences regarding reviewer effort. Finally, although the moderating effect of product type was not significant, follow-up analyses suggested that the nonlinear effect of arousal was stronger for utilitarian products, consistent with Study 1.

Despite the value of its experimental approach for establishing causation, conclusions from Study 3 are constrained by the fact that verbal review content was not identical across levels of arousal. Specifically, the three versions contained different emotional words (e.g., “happy” vs. “thrilled”), and reviews in the low-arousal condition contained one fewer sentence than those in the moderate- and high-arousal conditions. It is plausible that such differences influenced perceptions of either effort or helpfulness. Our next experiment addressed these concerns by manipulating arousal in a different manner, which relied exclusively on non-verbal cues.

STUDY 4

The design and procedure of the final study were similar to that of Study 3. However, stimuli were constructed in a manner by which arousal cues could be manipulated while holding constant the verbal content of reviews.

Stimulus Materials

Stimuli consisted of reviews for the same location-based reference apps presented in Study 3. We constructed the reviews using an approach similar to the prior study, with the following exceptions. First, we included only positive reviews.³ Second, we manipulated expressed arousal through the use of grammatical markers alone, without varying any words in the reviews. We applied two forms of grammatical markers: exclamation marks and capitalization (Allen 1988; Schandorf 2013). Reviews in the low-, moderate-, and high-arousal conditions were embedded with zero, two, and eight exclamation marks, respectively, and reviews in the high-arousal conditions contained five words written in capital letters. Third, in addition to the treatment versions of each review, we created a ‘baseline’ version as a control. Baseline reviews contained the descriptive sentences without the added emotional sentence, and without any exclamation marks or capitalization. Table 7 presents the four sets of reviews that resulted from this process.

- Insert Table 7 Here -

Procedure

One hundred fifty-seven undergraduate students participated in the study for course credit. The procedure was similar to that of Study 3. After seeing the cover story (Web Appendix E), participants read and evaluated five reviews, including one filler review followed by four treatment reviews. Treatment reviews were selected randomly from the four sets in Table 7, subject to the constraint that each review came from a different set, and the reviews were presented in random order. Measures of perceived helpfulness, arousal, and reviewer effort were identical to those in Study 3.

Results

Examination of the manipulation check confirmed that perceived arousal was greater in the high arousal reviews than that in the moderate arousal reviews ($M = 8.38$ vs. 6.85 , $t(156) = 17.62$, $p < .001$), which was in turn greater than that in the low arousal reviews ($M = 4.78$, $t(156) = 12.42$, $p < .001$). Therefore, the manipulation of expressed arousal was deemed successful. Perceived arousal in the baseline reviews ($M = 4.08$) was lower than that of all three treatment conditions (vs. low-arousal: $t(156) = -4.91$, $p < .001$).

To examine the effect of arousal on perceived helpfulness, we conducted an ANCOVA in which arousal was entered as a within-subject factor and review order was entered as a covariate. As before, the quadratic effect of arousal was specified as a polynomial contrast. Results revealed that the linear effect of arousal was negative and significant ($F(1, 133) = 35.856$, $p < .001$), indicating that reviews containing more arousal were on average deemed less helpful. More importantly, results revealed a significant quadratic effect of arousal ($F(1, 133) = 14.008$, $p < .001$). As illustrated in Figure 6, pairwise comparisons revealed no significant difference in perceived helpfulness from low to moderate arousal ($M = 6.61$ vs. 6.56 , $p > .7$), but a significant decrease in perceived helpfulness from moderate to high arousal ($M = 6.56$ vs. 5.53 , $t(156) = 6.29$, $p < .001$). Additional analyses indicated that reviews exhibiting low and medium levels of arousal were perceived as more helpful than baseline reviews ($M = 6.10$, $ts(156) = 2.96$ and 2.64 , $ps < .01$), but reviews exhibiting high levels of arousal were perceived as less helpful than baseline reviews ($t(156) = -2.87$, $p < .01$).

- Insert Figure 6 Here -

As in Study 3, we next investigated the mediating role of perceived reviewer effort by adopting the Judd et al. (2001) procedure for testing linear mediation in within-subject designs. Because helpfulness did not significantly differ between low and moderate levels of arousal, we

restricted the analysis to moderate and high levels. The first step revealed a negative effect of arousal on perceived effort ($M = 4.98$ vs. 4.56 , $t(156) = -3.23$, $p < .01$). The second step revealed that greater perceived effort was associated with greater helpfulness at both moderate and high levels of arousal ($\beta_s = .49$ and $.67$, $t_s = 7.01$ and 10.44 , $p_s < .001$). The third step revealed that differences in perceived effort predicted differences in helpfulness ($\beta = .63$, $t = 7.90$, $p < .001$). Thus, all three conditions were established, suggesting that the negative effect of arousal beyond moderate levels on perceived helpfulness was mediated by perceived effort.

Discussion

Supporting our main prediction, results of Study 4 again revealed a nonlinear effect of expressed arousal on perceived review helpfulness, even when the verbal content of reviews was held constant. Moreover, follow-up analyses confirmed that the detrimental effects of arousal at high levels could be partly explained by inferences regarding reviewer effort. In contrast to Studies 1-3, results of Study 4 did not reveal evidence for a beneficial effect of arousal at lower levels. Although unanticipated, this difference may simply reflect differing calibration across the four studies (see next section).

GENERAL DISCUSSION

Supplementing an emerging stream of research on the perceived value of WOM for decision-making (Chen and Lurie 2013; He and Bond 2013; Moore 2015; Mudambi and Schuff 2010; Yin, Bond, and Zhang 2014), our research explored the consequences of emotional arousal in consumer reviews on subsequent reader perceptions. Four studies offered methodological triangulation through archival data analysis, surveys, and laboratory experiments. Expressed arousal was both measured and manipulated across our studies, using verbal cues (Study 1), non-

verbal cues (Study 4), and a combination (Studies 2 and 3). Table 8 summarizes major findings. Supporting our main prediction, the four studies provided consistent evidence that expressed arousal affects reader perceptions of review helpfulness in a nonlinear manner of diminishing returns. In addition, we obtained both direct and indirect support for an inference-based account, in which the nonlinear effect of arousal operates through reader inferences regarding the effort expended by the reviewer in constructing his or her review. Evidence for the moderating role of product type was not conclusive; however, the pattern of results across studies was consistent with our claim that that the nonlinear effect of arousal is magnified for utilitarian products.

- Insert Table 8 Here -

Theoretical Implications

Our findings contribute to growing scholarship on consumer WOM and decision-making, as well as the broader topic of affective processes in communication. Existing research on determinants of review ‘helpfulness’ has tended to focus on readily observable variables, including ratings, reviewer characteristics, etc. (Chen and Lurie 2013; Forman, Ghose, and Wiesenfeld 2008; Mudambi and Schuff 2010; Yin, Mitra, and Zhang 2016). However, an emerging trend has begun to explore review content more directly (Cao, Duan, and Gan 2011; Moore 2015; Yin, Bond, and Zhang 2014), and we supplement this trend by investigating the role of expressed emotion, which is abundant in review settings but has received little prior attention.

Almost all prominent emotion frameworks include fundamental dimensions of valence and arousal (Niedenthal 2008; Russell 1980). Within the vast marketing scholarship on affect and emotion, research examining the valence dimension is plentiful, but research on the arousal dimension is rare (c.f. Fedorikhin and Patrick 2010; Gorn, Pham, and Sin 2001). Reflecting this

pattern, consumer WOM research has tended to focus on valence (Chen and Lurie 2013; Ludwig et al. 2013), but a few notable exceptions have explored other aspects. Yin, Bond, and Zhang (2014) examined the influence of discrete emotions (such as anxiety and anger) in consumer reviews, demonstrating that same-valence emotions often have distinct effects on reader perceptions. The present findings suggest that over and above the specific emotions conveyed, review-embedded arousal cues themselves directly influence inferences regarding both reviewer and review. Berger and colleagues have incorporated arousal directly in a model of WOM transmission (Berger 2011; Berger and Milkman 2012), showing that message content which evokes greater arousal in perceivers is more likely to be shared by recipients. Our research focuses instead on the arousal of senders, as signaled by cues in their messages, and is consistent with an ‘actor-observer’ framework in which recipients use evidence of arousal to inform message-relevant judgments.

Research on affect and information processing provides a complex picture regarding the impact of emotional arousal on cognitive performance and task outcomes (Naqvi, Shiv, and Bechara 2006; Tice, Bratslavsky, and Baumeister 2001), and well-known findings in the marketing domain have linked elevated arousal to both positive and negative consumer outcomes (e.g., Isen 2001; Shiv and Fedorikhin 1999). Moreover, consumers differ in the extent to which they view affective experience as detrimental or beneficial to rational judgment (Avnet, Pham, and Stephen 2012; Epstein 1994; Hsee et al. 2003; Isen 2001). Acknowledging this ambivalence, our framework allows for a nuanced role of emotional arousal in interpersonal communication, in which the expression of additional arousal can have beneficial, neutral, or even detrimental effect on the perceived value of message content, depending on the level of pre-existing arousal and the situational context. More generally, our findings support a growing consensus that

reliance on affect is heavily dependent on situational factors, even for the same individual and judgment (Pham 2009).

At a broader level, our inference-based account reflects a growing trend in which scholars are beginning to explore the various interpersonal, communicative functions served by emotion (Hareli and Hess 2012). Emotional experiences not only motivate and guide our own behavior, but also provide observers with valuable social information (Van Kleef, De Dreu, and Manstead 2010). The social functions of emotion may be especially relevant in the realm of online communication, where messages have a long ‘shelf-life’ and recipients tend to have little information about the sender or context. Our findings shed light on the process by which lay theories are applied by consumers to process the information conveyed by review-embedded emotion.

Implications for WOM Collection and Management

Many real-world review platforms offer formal guidelines for prospective reviewers. These guidelines vary considerably with respect to the expression of emotion: some platforms advise review authors to restrain their feelings, while others recommend that authors express those feelings freely (e.g., the Yelp and Amazon examples presented earlier). Our findings suggest that in many contexts, neither approach is optimal: reviewers who appear overly ‘relaxed’ or overly ‘fired up’ may be perceived as less helpful by future readers, even when they have provided objectively useful and relevant information. Instead, an implication of our research is that guidelines be designed to encourage a moderate level of emotional expression (e.g., “express your feelings freely, without holding back or exaggerating.”). Moreover, our findings suggest that appropriate instructions may vary systematically across settings. In particular, guidelines

concerning emotional expression may be especially useful for retailers in utilitarian categories, where the effects of arousal on perceived helpfulness appear to be strongest.

Marketers are increasingly turning to online forums as a tool for customer research and relationship management (Decker and Trusov 2010; Ludwig et al. 2013). Faced with an enormous volume of communications, firms often wish to identify those messages (and senders) which are especially ‘influential.’ However, no definitive measure of message influence exists, and available proxies (comments, likes, helpfulness votes, etc.) require time to accumulate. Complementing research on characteristics of viral WOM (Berger and Milkman 2012; Chen and Lurie 2013), our work offers implications for marketers seeking to develop methods of estimating, a priori, the extent to which specific WOM will be deemed helpful. In addition to other relevant variables (source characteristics, information content, etc.), our findings suggest that message-embedded arousal should be considered directly in this estimation. Using readily available software tools (e.g., the RDAL applied in Study 1), measurement of expressed arousal can be automated and incorporated into existing communication monitoring tools.

Limitations and Future Research

Although findings of all four studies supported the pattern of diminishing returns predicted by H1, the specific form of the observed relationship differed. Such variance is unsurprising given the different contexts and stimuli used across the studies, and may simply result from differences in calibration (i.e., arousal that seems ‘extreme’ in one context may seem ‘moderate’ in other contexts). In fact, the inflection point of the observed nonlinear curve varied substantially across studies. Following this logic, the lack of a positive effect of arousal at lower levels in Study 4 may reflect the fact that “low-arousal” reviews were close to the inflection point of the nonlinear curve. Supporting this speculation, a comparison of manipulation checks

suggests that perceived arousal in the low-arousal condition was notably higher in Study 4 than Study 3 ($M = 4.78$ vs. 4.14). More generally, perceptions of extreme arousal may be difficult to induce in experimental settings. Nonetheless, our findings suggest an opportunity to identify specific contexts under which different patterns of nonlinearity obtain.

Our theoretical arguments (and results of Studies 3 and 4) suggest that the nonlinear effect of expressed arousal can be explained in part by inferences regarding reviewer effort, which are known to play a critical role in helpfulness perceptions (Mudambi and Schuff 2010; Yin, Bond, and Zhang 2014). However, we acknowledge that a wide variety of other mechanisms may contribute to the nonlinear effect. For instance, reviews exhibiting a moderate level of arousal might be processed more fluently, resulting in more favorable global evaluations (e.g., Lee and Labroo 2004). Readers may perceive high-arousal reviewers to be more subjective rather than objective (e.g., 'gushing' or 'venting'; Jensen et al. 2013; Park, Lee, and Han 2007), or readers may assume that such reviewers are “trying too hard” to influence prospective consumers, undermining their own credibility (Friestad and Wright 1994; Xu and Wyer Jr 2010).⁴ The interplay of these factors is worthy of future investigation.

Similarly, our arguments for the moderating role of product type (H3) rested on the assumption that emotional cues will be deemed more diagnostic when they are unexpected or atypical. This logic is consistent with the view that emotions which “stand out” are more likely to be used in decision making (Albarracín and Kumkale 2003; Greifeneder, Bless, and Pham 2011; Siemer and Reizenzein 1998). A wide range of other contextual factors are likely to alter readers' emotional expectations: reader and reviewer demographics, the ‘tone’ of other reviews on the platform, etc. We encourage systematic exploration of such factors, and in particular their influence on the relationship between expressed arousal and reader inference.

Our investigation focused solely on the perceived usefulness of consumer reviews, as judged by their readers. Thus it would be worthwhile to examine other downstream variables. For example, how does the arousal expressed in a product review affect its persuasiveness, subsequent decision confidence, etc.? Are reviews containing more arousal more impactful (even if they are not deemed helpful)? Are they more likely to be elaborated on, encoded, remembered accurately, etc.? Our findings are clearly relevant, but the questions merit additional research.

Our findings extend understanding of the process by which WOM is interpreted, and they highlight the importance of incorporating emotional arousal into existing communication frameworks. Although our studies were conducted solely in the context of online reviews, we expect our main arguments to generalize to other WOM settings (online communities, social media platforms, etc.). However, more research is needed to test the robustness of our findings across settings and identify potential contingencies. Different communication channels are likely to evoke different consumer expectations regarding arousal or effort (see above): compared to a product review, for example, readers may expect higher levels of arousal in a product-related Facebook post, personal blog entry, etc. Similarly, forums differ markedly in the extent to which the identity of authors is disclosed. Recent findings in online WOM document an egocentric process, in which message recipients tend to assume underlying similarity with senders unless explicitly informed otherwise (He and Bond 2013; Naylor, Lamberton, and Norton 2011). It would be worthwhile to explore the interplay of expressed arousal, identity disclosure, and assumed similarity in shaping helpfulness perceptions.

REFERENCES

- Albarracín, Dolores and G. Tarcan Kumkale (2003), "Affect as Information in Persuasion: A Model of Affect Identification and Discounting," *Journal of Personality and Social Psychology*, 84 (3), 453-69.
- Allen, Thomas B. (1988), "Bulletin Boards of the 21st Century Are Coming of Age," *Smithsonian*, 19 (6), 83-93.
- Amazon (2016), "About Customer Reviews," (accessed Mar. 12, 2016), [available at <http://www.amazon.com/gp/help/customer/display.html?ie=UTF8&nodeId=201077870>].
- Avnet, Tamar, Michel Tuan Pham, and Andrew T. Stephen (2012), "Consumers' Trust in Feelings as Information," *Journal of Consumer Research*, 39 (4), 720-35.
- Babin, Barry J., William R. Darden, and Mitch Griffin (1994), "Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value," *Journal of Consumer Research*, 20 (4), 644-56.
- Bachorowski, Jo-Anne and Ellen B. Braaten (1994), "Emotional Intensity: Measurement and Theoretical Implications," *Personality and Individual Differences*, 17 (2), 191-99.
- Baum, Christopher F. (2008), "Stata Tip 63: Modeling Proportions," *Stata Journal*, 8 (2), 299-303.
- Berger, Jonah (2011), "Arousal Increases Social Transmission of Information," *Psychological Science*, 22 (7), 891-93.
- Berger, Jonah and Katherine L. Milkman (2012), "What Makes Online Content Viral?," *Journal of Marketing Research*, 49 (2), 192-205.
- Botti, Simona and Ann L. McGill (2011), "The Locus of Choice: Personal Causality and Satisfaction with Hedonic and Utilitarian Decisions," *Journal of Consumer Research*, 37 (6), 1065-78.
- Brosch, Tobias, Gilles Pourtois, and David Sander (2010), "The Perception and Categorisation of Emotional Stimuli: A Review," *Cognition & Emotion*, 24 (3), 377-400.
- Cannon, Walter B. (1927), "The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory," *The American Journal of Psychology*, 39 (1/4), 106-24.
- Cao, Qing, Wenjing Duan, and Qiwei Gan (2011), "Exploring Determinants of Voting for the "Helpfulness" of Online User Reviews: A Text Mining Approach," *Decision Support Systems*, 50 (2), 511-21.
- Carey, John (1980), "Paralanguage in Computer Mediated Communication," in Proceedings of the 18th annual meeting on Association for Computational Linguistics. Philadelphia, Pennsylvania: Association for Computational Linguistics.
- Chen, Yubo and Jinhong Xie (2008), "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science*, 54 (3), 477-91.
- Chen, Zoey and Nicholas H. Lurie (2013), "Temporal Contiguity and Negativity Bias in the Impact of Online Word-of-Mouth," *Journal of Marketing Research*, 50 (4), 463-76.
- Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (3), 345-54.
- Damasio, Antonio (2005), *Descartes' Error: Emotion, Reason, and the Human Brain*: Penguin.
- Decker, Reinhold and Michael Trusov (2010), "Estimating Aggregate Consumer Preferences from Online Product Reviews," *International Journal of Research in Marketing*, 27 (4), 293-307.
- Dellarocas, Chrysanthos (2003), "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science*, 49 (10), 1407-24.

- Dhar, Ravi and Klaus Wertenbroch (2000), "Consumer Choice between Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 37 (1), 60-71.
- Ekman, Paul (1993), "Facial Expression and Emotion," *American Psychologist*, 48 (4), 384-92.
- Epstein, Seymour (1994), "Integration of the Cognitive and the Psychodynamic Unconscious," *American Psychologist*, 49 (8), 709-24.
- Fedorikhin, Alexander and Vanessa M. Patrick (2010), "Positive Mood and Resistance to Temptation: The Interfering Influence of Elevated Arousal," *Journal of Consumer Research*, 37 (4), 698-711.
- Forman, Chris, Anindya Ghose, and Batia Wiesenfeld (2008), "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research*, 19 (3), 291-313.
- Friestad, Marian and Peter Wright (1994), "The Persuasion Knowledge Model: How People Cope with Persuasion Attempts," *Journal of Consumer Research*, 21 (1), 1-31.
- Frijda, Nico H., Peter Kuipers, and Elisabeth Ter Schure (1989), "Relations among Emotion, Appraisal, and Emotional Action Readiness," *Journal of Personality and Social Psychology*, 57 (2), 212-28.
- Fritz, Matthew S. and David P. Mackinnon (2007), "Required Sample Size to Detect the Mediated Effect," *Psychological Science*, 18 (3), 233-39.
- Gendron, Maria, Kristen A. Lindquist, Lawrence Barsalou, and Lisa Feldman Barrett (2012), "Emotion Words Shape Emotion Percepts," *Emotion*, 12 (2), 314-25.
- Gernsbacher, Morton Ann, Brenda M. Hallada, and Rachel R. W. Robertson (1998), "How Automatically Do Readers Infer Fictional Characters' Emotional States?," *Scientific Studies of Reading*, 2 (3), 271-300.
- Gershoff, Andrew D., Susan M. Broniarczyk, and Patricia M. West (2001), "Recommendation or Evaluation? Task Sensitivity in Information Source Selection," *Journal of Consumer Research*, 28 (3), 418-38.
- Gorn, Gorn, Michel Tuan Pham, and Leo Yatming Sin (2001), "When Arousal Influences Ad Evaluation and Valence Does Not (and Vice Versa)," *Journal of Consumer Psychology*, 11 (1), 43-55.
- Greifeneder, Rainer, Herbert Bless, and Michel Tuan Pham (2011), "When Do People Rely on Affective and Cognitive Feelings in Judgment? A Review," *Personality and Social Psychology Review*, 15 (2), 107-41.
- Gunning, Robert (1969), "The Fog Index after Twenty Years," *Journal of Business Communication*, 6 (2), 3-13.
- Hareli, Shlomo and Ursula Hess (2012), "The Social Signal Value of Emotions," *Cognition & Emotion*, 26 (3), 385-89.
- Harris, Ranida B and David Paradise (2007), "An Investigation of the Computer-Mediated Communication of Emotions," *Journal of Applied Sciences Research*, 3 (12), 2081-90.
- Hayes, Andrew F. and Kristopher J. Preacher (2010), "Quantifying and Testing Indirect Effects in Simple Mediation Models When the Constituent Paths Are Nonlinear," *Multivariate Behavioral Research*, 45 (4), 627-60.
- He, Stephen and Samuel D. Bond (2013), "Word-of-Mouth and the Forecasting of Consumption Enjoyment," *Journal of Consumer Psychology*, 23 (4), 464-82.
- Heilman, Kenneth M. (1997), "The Neurobiology of Emotional Experience," *Journal of Neuropsychiatry and Clinical Neurosciences*, 9 (3), 439-48.

- Hsee, Christopher K., Jiao Zhang, Fang Yu, and Yiheng Xi (2003), "Lay Rationalism and Inconsistency between Predicted Experience and Decision," *Journal of Behavioral Decision Making*, 16 (4), 257-72.
- Huddy, Leonie, Stanley Feldman, and Erin Cassese (2007), "On the Distinct Political Effects of Anxiety and Anger," in *The Affect Effect: Dynamics of Emotion in Political Thinking and Behavior*, Ann Crigler and Michael Mackuen and George E. Marcus and W. Russell Neuman, eds.: University of Chicago Press.
- Isen, Alice M. (2001), "An Influence of Positive Affect on Decision Making in Complex Situations: Theoretical Issues with Practical Implications," *Journal of Consumer Psychology*, 11 (2), 75-85.
- Jensen, Matthew L., Joshua M. Averbach, Zhu Zhang, and Kevin B. Wright (2013), "Credibility of Anonymous Online Product Reviews: A Language Expectancy Perspective," *Journal of Management Information Systems*, 30 (1), 293-324.
- Johnson, Eric J. and Amos Tversky (1983), "Affect, Generalization, and the Perception of Risk," *Journal of Personality and Social Psychology*, 45 (1), 20-31.
- Judd, Charles M., David A. Kenny, and Gary H. McClelland (2001), "Estimating and Testing Mediation and Moderation in within-Subject Designs," *Psychological Methods*, 6 (2), 115-34.
- Korfiatis, Nikos, Daniel Rodriguez, and M. Angel Sicilia (2008), "The Impact of Readability on the Usefulness of Online Product Reviews: A Case Study on an Online Bookstore," in *Emerging Technologies and Information Systems for the Knowledge Society, Proceedings*, Miltiadis D. Lytras and John M. Carroll and Ernesto Damiani and Robert D. Tennyson, eds. Vol. 5288. Berlin: Springer-Verlag Berlin.
- Kronrod, Ann and Shai Danziger (2013), "'Wii Will Rock You!' The Use and Effect of Figurative Language in Consumer Reviews of Hedonic and Utilitarian Consumption," *Journal of Consumer Research*, 40 (4), 726 - 39.
- Lang, Peter J. (1995), "The Emotion Probe - Studies of Motivation and Attention," *American Psychologist*, 50 (5), 372-85.
- Lee, Angela Y. and Aparna A. Labroo (2004), "The Effect of Conceptual and Perceptual Fluency on Brand Evaluation," *Journal of Marketing Research*, 41 (2), 151-65.
- Lerner, Jennifer S. and Dacher Keltner (2000), "Beyond Valence: Toward a Model of Emotion-Specific Influences on Judgement and Choice," *Cognition & Emotion*, 14 (4), 473-93.
- Levenson, Robert W (1994), "Human Emotion: A Functional View," in *The Nature of Emotion: Fundamental Questions*, Paul Ekman and Richard J. Davidson, eds. New York: Oxford University Press.
- Loewenstein, George (1996), "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, 65 (3), 272-92.
- Ludwig, Stephan, Ko De Ruyter, Mike Friedman, Elisabeth C. Brügger, Martin Wetzels, and Gerard Pfann (2013), "More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates," *Journal of Marketing*, 77 (1), 87-103.
- Mano, Haim (1991), "The Structure and Intensity of Emotional Experiences - Method and Context Convergence," *Multivariate Behavioral Research*, 26 (3), 389-411.
- Moore, Sarah G. (2015), "Attitude Predictability and Helpfulness in Online Reviews: The Role of Explained Actions and Reactions," *Journal of Consumer Research*, 42 (1), 30-44.

- Mudambi, Susan M. and David Schuff (2010), "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.Com," *MIS Quarterly*, 34 (1), 185-200.
- Naqvi, Nasir, Baba Shiv, and Antoine Bechara (2006), "The Role of Emotion in Decision Making: A Cognitive Neuroscience Perspective," *Current Directions in Psychological Science*, 15 (5), 260-64.
- Naylor, Rebecca Walker, Cait Poyner Lamberton, and David A. Norton (2011), "Seeing Ourselves in Others: Reviewer Ambiguity, Egocentric Anchoring, and Persuasion," *Journal of Marketing Research*, 48 (3), 617-31.
- Niedenthal, Paula M. (2008), "Emotion Concepts," in *Handbook of Emotions*, Michael Lewis and Jeannette M. Haviland-Jones and Lisa Feldman Barrett, eds. 3 ed. New York: The Guilford Press.
- O'hanlon, James F. (1981), "Boredom - Practical Consequences and a Theory," *Acta Psychologica*, 49 (1), 53-82.
- Oatley, Keith and P. N. Johnson-Laird (2014), "Cognitive Approaches to Emotions," *Trends in Cognitive Sciences*, 18 (3), 134-40.
- Osgood, Charles E. (1966), "Dimensionality of the Semantic Space for Communication Via Facial Expressions," *Scandinavian Journal of Psychology*, 7 (1), 1-30.
- Park, Do-Hyung, Jumin Lee, and Ingoo Han (2007), "The Effect of on-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement," *International Journal of Electronic Commerce*, 11 (4), 125-48.
- Pham, Michel Tuan (2007), "Emotion and Rationality: A Critical Review and Interpretation of Empirical Evidence," *Review of General Psychology*, 11 (2), 155-78.
- (2009), "The Lexicon and Grammar of Affect as Information in Consumer Decision-Making: The Gaim," in *Social Psychology of Consumer Behavior*, Michaela Wänke, ed. New York: Psychology Press.
- Planalp, Sally (1998), "Communicating Emotion in Everyday Life: Cues, Channels, and Processes," in *Handbook of Communication and Emotion: Research, Theory, Applications, and Contexts*, Peter A. Andersen and Laura K. Guerrero, eds. San Diego, CA: Academic Press.
- Russell, James A. (1980), "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, 39 (6), 1161-78.
- Russell, James A. and Lisa Feldman Barrett (1999), "Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant," *Journal of Personality and Social Psychology*, 76 (5), 805-19.
- Sanbonmatsu, David M. and Frank R. Kardes (1988), "The Effects of Physiological Arousal on Information-Processing and Persuasion," *Journal of Consumer Research*, 15 (3), 379-85.
- Schachter, Stanley and Jerome Singer (1962), "Cognitive, Social, and Physiological Determinants of Emotional State," *Psychological Review*, 69 (5), 379-99.
- Schandorf, Michael (2013), "Mediated Gesture: Paralinguistic Communication and Phatic Text," *Convergence: The International Journal of Research into New Media Technologies*, 19 (3), 319-44.
- Scherer, Klaus R. (2000), "Emotion," in *Introduction to Social Psychology: A European Perspective*, Miles Hewstone and Wolfgang Stroebe, eds. 3 ed. Oxford: Blackwell.
- Schindler, Robert M. and Barbara Bickart (2012), "Perceived Helpfulness of Online Consumer Reviews: The Role of Message Content and Style," *Journal of Consumer Behaviour*, 11 (3), 234-43.

- Sen, Shahana and Dawn Lerman (2007), "Why Are You Telling Me This? An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing*, 21 (4), 76-94.
- Seo, Myeong-Gu, Lisa Feldman Barrett, and Jean M. Bartunek (2004), "The Role of Affective Experience in Work Motivation," *Academy of Management Review*, 29 (3), 423-39.
- Shaver, Phillip, Judith Schwartz, Donald Kirson, and Cary O'Connor (1987), "Emotion Knowledge: Further Exploration of a Prototype Approach," *Journal of Personality and Social Psychology*, 52 (6), 1061-86.
- Shiv, Baba and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of Consumer Research*, 26 (3), 278-92.
- Siemer, Matthias and Rainer Reisenzein (1998), "Effects of Mood on Evaluative Judgements: Influence of Reduced Processing Capacity and Mood Salience," *Cognition & Emotion*, 12 (6), 783-805.
- Skinner, Ellen A., Michael Chapman, and Paul B. Baltes (1988), "Control, Means-Ends, and Agency Beliefs: A New Conceptualization and Its Measurement During Childhood," *Journal of Personality and Social Psychology*, 54 (1), 117-33.
- Sloot, Laurens M., Peter C. Verhoef, and Philip Hans Franses (2005), "The Impact of Brand Equity and the Hedonic Level of Products on Consumer Stock-out Reactions," *Journal of Retailing*, 81 (1), 15-34.
- Smith, Craig A. and Phoebe C. Ellsworth (1985), "Patterns of Cognitive Appraisal in Emotion," *Journal of Personality and Social Psychology*, 48 (4), 813-38.
- Tice, Dianne M., Ellen Bratslavsky, and Roy F. Baumeister (2001), "Emotional Distress Regulation Takes Precedence over Impulse Control: If You Feel Bad, Do It!," *Journal of Personality and Social Psychology*, 80 (1), 53-67.
- Tormala, Zakary L., Joshua J. Clarkson, and Marlone D. Henderson (2011), "Does Fast or Slow Evaluation Foster Greater Certainty?," *Personality and Social Psychology Bulletin*, 37 (3), 422-34.
- Van Kleef, Gerben A. (2010), "The Emerging View of Emotion as Social Information," *Social and Personality Psychology Compass*, 4 (5), 331-43.
- Van Kleef, Gerben A., Carsten K. W. De Dreu, and Antony S. R. Manstead (2010), "An Interpersonal Approach to Emotion in Social Decision Making: The Emotions as Social Information Model," in *Advances in Experimental Social Psychology*, Vol 42, Vol. 42. San Diego: Elsevier Academic Press Inc.
- Vandergriff, Ilona (2013), "Emotive Communication Online: A Contextual Analysis of Computer-Mediated Communication (CMC) Cues," *Journal of Pragmatics*, 51 (0), 1-12.
- Vygotsky, Lev S. (1964), "Thought and Language," *Annals of Dyslexia*, 14 (1), 97-98.
- Watson, David and Auke Tellegen (1985), "Toward a Consensual Structure of Mood," *Psychological Bulletin*, 98 (2), 219-35.
- Weiner, Bernard and Andy Kukla (1970), "An Attributional Analysis of Achievement Motivation," *Journal of Personality and Social Psychology*, 15 (1), 1-20.
- Whissell, Cynthia (2009), "Using the Revised Dictionary of Affect in Language to Quantify the Emotional Undertones of Samples of Natural Language," *Psychological Reports*, 105 (2), 509-21.
- Whissell, Cynthia, Michael Fournier, Rene Pelland, Deborah Weir, and K. Makarec (1986), "A Dictionary of Affect in Language: Iv. Reliability, Validity, and Applications," *Perceptual and Motor Skills*, 62 (3), 875-88.

- Xu, Alison Jing and Robert S. Wyer Jr (2010), "Puffery in Advertisements: The Effects of Media Context, Communication Norms, and Consumer Knowledge," *Journal of Consumer Research*, 37 (2), 329-43.
- Yelp (2013), "How Yelp Helps You Find the Right Local Business," (accessed Nov. 13, 2014), [available at <http://officialblog.yelp.com/2013/11/yelp-recommended-reviews.html>].
- Yin, Dezhi, Samuel D. Bond, and Han Zhang (2014), "Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews," *MIS Quarterly*, 38 (2), 539-60.
- Yin, Dezhi, Sabyasachi Mitra, and Han Zhang (2016), "When Do Consumers Value Positive Vs. Negative Reviews? An Empirical Investigation of Confirmation Bias in Online Word of Mouth," *Information Systems Research*, 27 (1), 131-44.

FOOTNOTES

¹ As a robustness check, we followed prior research (Mudambi and Schuff 2010) by measuring review helpfulness as the proportion of helpful votes out of total votes. Analyses utilized fractional logit models to account for the bounded dependent variable (Baum 2008), and the sample was restricted to reviews with at least five votes. Results were nearly identical to those of the main analysis, and all significant tests reported below remained significant.

² We conducted a robustness check by including squared terms for both review rating and length in our main analysis. All significant results reported here remained significant (see Web Appendix B).

³ In Studies 1-2, the estimated coefficient for the interaction of review rating and squared arousal was positive and significant ($p < .05$), suggesting that the nonlinear impact of arousal was greater for negative reviews than positive reviews. Given that valence is not part of our theoretical framework, we decided to hold valence constant in Study 4.

⁴ A common theme among many of these alternative inferences is that the reviewer may have been thinking irrationally. Measures of perceived reviewer rationality were collected in Studies 2-4, and supplementary analyses are presented in Web Appendix F. Results suggest that at high levels of arousal, the negative effects of additional arousal are explained in part by negative inferences about reviewer rationality.

TABLES

Table 1 – Study 1: Descriptive Statistics and Variable Correlations (N = 414,336)

Variable	Mean	S.D.	Min	Max	1	2	3	4	5	6	7
1 Helpful votes	2.37	7.10	0	984	1						
2 Total votes	3.87	9.87	1	1031	0.93	1					
3 Rating	3.45	1.68	1	5	0.05	-0.03	1				
4 Length	42.02	49.04	1	1134	0.12	0.09	0.03	1			
5 Reading difficulty	7.09	4.07	0.4	461.6	0.04	0.03	0.04	0.33	1		
6 Emotional valence	1.89	0.14	1	3	0.00	-0.02	0.33	-0.13	-0.11	1	
7 Emotional arousal	1.68	0.12	1	2.89	-0.02	-0.02	0.10	-0.16	-0.15	0.38	1

Notes: The operationalization of each variable is described in Table 2.

Table 2 – Study 1: Variables and Operationalizations

Variable Type	Variable Level	#	Variable	Operationalization	Notes
DV	Individual Review	1	Helpful Votes	# of helpful votes	
IV	Individual Review	2	Emotional Arousal	average arousal score of identified words	Range: [1, 3] Coded by Revised Dictionary of Affect in Language (RDAL)
Moderator	Product Category	3	Utilitarian value	average utilitarian value of the app category	Range: [-8, 8]
		4	Total Votes	# of total votes	
		5	Rating	# of stars	Range: [1, 5]
		6	Length	# of words	
Control	Individual Review	7	Reading difficulty	Gunning Fog Index	Number of years of formal education needed to understand the text on a first reading
		8	Emotional valence	average valence score of identified words	Also coded by RDAL
	Product	9	Quality	average rating	Range: [1, 5]
		10	Popularity	# of ratings in total	
		11	Price	=1 if the app is paid	

Table 3 – Study 1: Negative Binomial Model Coefficients

DV: number of helpful votes	Model 1	Model 2
Number of Total Votes	0.109*** (0.001)	0.109*** (0.001)
Rating	0.154*** (0.001)	0.154*** (0.001)
Length	0.002*** (0.000)	0.002*** (0.000)
Reading difficulty	0.008*** (0.001)	0.008*** (0.001)
Average rating	-0.092*** (0.002)	-0.092*** (0.002)
Number of ratings	-0.000*** (0.000)	-0.000*** (0.000)
Paid	0.099*** (0.003)	0.099*** (0.003)
Utilitarian value	0.006*** (0.000)	-0.054* (0.032)
Emotional valence	0.083*** (0.012)	0.086*** (0.012)
Emotional arousal	2.979*** (0.185)	3.243*** (0.197)
Arousal ²	-0.883*** (0.053)	-0.966*** (0.057)
Arousal × Utilitarian value		0.081** (0.037)
Arousal ² × Utilitarian value		-0.027** (0.010)
Constant	-2.986*** (0.164)	-3.201*** (0.174)
N	414336	414336
Log Likelihood	-681765.4	-681753.0

Notes: Standard errors in parentheses; *** significant at 0.01, ** significant at 0.05, * significant at 0.1

Table 4 – Study 2: Descriptive Statistics and Variable Correlations (N = 400)

Variable	Mean	S.D.	Min	Max	1	2	3	4	5	6	7
1 Review helpfulness	5.42	1.76	1	8.7	1						
2 Rating	3.65	1.59	1	5	-0.03	1					
3 Length	46.66	39.10	3	350	0.53	-0.03	1				
4 Reading difficulty	6.90	2.97	0.6	16.7	0.42	0.07	0.43	1			
5 Emotional valence	5.71	1.94	1	9	0.16	0.82	0.08	0.15	1		
6 Emotional arousal	5.76	1.30	1	8.4	0.20	0.12	0.24	0.04	0.22	1	
7 Effort	5.01	1.81	1	9	0.86	0.04	0.70	0.52	0.20	0.27	1

Table 5 – Study 2: Regression Model Coefficients

DV: review helpfulness	Model 1	Model 2
Rating	-0.334*** (0.077)	-0.333*** (0.077)
Length	0.017*** (0.002)	0.017*** (0.002)
Reading difficulty	0.116*** (0.026)	0.116*** (0.026)
Average rating	-0.130 (0.103)	-0.136 (0.103)
Number of ratings	0.000 (0.000)	0.000 (0.000)
Paid	0.032 (0.144)	0.034 (0.144)
Utilitarian value	0.043*** (0.015)	0.134 (0.208)
Emotional valence	0.310*** (0.064)	0.308*** (0.065)
Emotional arousal	1.414*** (0.351)	1.232*** (0.430)
Arousal ²	-0.120*** (0.032)	-0.101** (0.040)
Arousal × Utilitarian value		-0.048 (0.078)
Arousal ² × Utilitarian value		0.005 (0.007)
Constant	-0.199 (0.971)	0.234 (1.162)
N	400	400
Adjusted R ²	0.3959	0.3949

Notes: Standard errors in parentheses; *** significant at 0.01, ** significant at 0.05, * significant at 0.1

Table 6 – Study 3: Review Stimuli

Location App (Utilitarian)

#	Low Arousal	Moderate Arousal	High Arousal
1	Clean and fast. However, I noticed that it is not able to look up places other than your current location.	Very clean and fast! However, I was surprised that it is not able to look up places other than your current location! That was disappointing!	Very clean and fast!!! However, I was COMPLETELY shocked that it is not able to look up places other than your current location!!! That was really disappointing!!!
2	Waste of money. It does organize information into a graphical view, but you cannot enter a location of interest.	Waste of money! It does organize information into a graphical view, but it bothers me that you cannot enter a location of interest! I am not pleased!	Waste of money!!! It does organize information into a graphical view, but it REALLY annoys me that you cannot enter a location of interest!!! I am outraged!!!
3	High quality app. It tells you how far away you are from the landmarks and even shows those landmarks on a map.	High quality app! It tells you how far away you are from the landmarks and even shows those landmarks on a map! Happy with it!	High quality app!!! It tells you how far away you are from the landmarks and EVEN shows those landmarks on a map!!! Thrilled with it!!!

Fish Pond App (Hedonic)

#	Low Arousal	Moderate Arousal	High Arousal
1	Fun and cool. However, I noticed that the fish swim off too fast and easily when you tap them.	Very fun and cool! However, I was surprised that the fish swim off too fast and easily when you tap them! That was disappointing!	Very fun and cool!!! However, I was COMPLETELY shocked that the fish swim off too fast and easily when you tap them!!! That was really disappointing!!!
2	Waste of money. The pond environment is pretty, but the color scheme is muddy and uninteresting.	Waste of money! The pond environment is pretty, but it bothers me that the color scheme is muddy and uninteresting! I am not pleased!	Waste of money!!! The pond environment is pretty, but it REALLY annoys me that the color scheme is muddy and uninteresting!!! I am outraged!!!
3	Entertaining app. You can use your finger on the screen to interact with the fish, as if you were actually there.	Entertaining app! You can use your finger on the screen to interact with the fish, as if you were actually there! Happy with it!	Entertaining app!!! You can use your finger on the screen to interact with the fish, as if you were ACTUALLY there!!! Thrilled with it!!!

Table 7 – Study 4: Review Stimuli

#	Baseline	Low Arousal	Moderate Arousal	High Arousal
1	High quality app. You can enter a location of interest, and it is able to look up places other than your current location. This is convenient for trip planning.	High quality app. You can enter a location of interest, and it is able to look up places other than your current location. This is convenient for trip planning. I am pleased.	High quality app! You can enter a location of interest, and it is able to look up places other than your current location. This is convenient for trip planning. I am pleased!	HIGH QUALITY app!!! You can enter a location of interest, and it is able to look up places OTHER than your current location! This is CONVENIENT for trip planning! I am PLEASED!!!
2	The interface is easy to use. It provides a nice graphical view of the information. Plus, it can show landmarks on a map and tell you how far away they are.	The interface is easy to use. It provides a nice graphical view of the information. Plus, it can show landmarks on a map and tell you how far away they are. I'm very happy with this.	The interface is easy to use! It provides a nice graphical view of the information. Plus, it can show landmarks on a map and tell you how far away they are. I'm very happy with this!	The interface is EASY to use!!! It provides a NICE graphical view of the information! Plus, it can show landmarks on a map AND tell you how far away they are! I'm VERY HAPPY with this!!!
3	Lots of options and customizable. The text size is readable and can be adjusted. And it allows white text on black background, which is much easier on the eyes.	I love it. Lots of options and customizable. The text size is readable and can be adjusted. And it allows white text on black background, which is much easier on the eyes.	I love it! Lots of options and customizable. The text size is readable and can be adjusted. And it allows white text on black background, which is much easier on the eyes!	I LOVE it!!! Lots of options and CUSTOMIZABLE! The text size is READABLE and can be adjusted! And it allows white text on black background, which is MUCH EASIER on the eyes!!!
4	Reading articles is like flipping through a real book. It makes the information easy to access. When you tap a topic, related topics spread out from it in a nice animation.	I really like this app. Reading articles is like flipping through a real book. It makes the information easy to access. When you tap a topic, related topics spread out from it in a nice animation.	I really like this app! Reading articles is like flipping through a real book. It makes the information easy to access. When you tap a topic, related topics spread out from it in a nice animation!	I REALLY LIKE this app!!! Reading articles is like flipping through a REAL book! It makes the information EASY to access! When you tap a topic, related topics spread out from it in a NICE animation!!!

Table 8 – Summary of Findings

Hypotheses	Study 1	Study 2	Study 3	Study 4
H1: The effect of arousal on review helpfulness follows a trend of diminishing returns	Supported (inverted U-shaped)	Supported (inverted U-shaped)	Supported (diminishing returns)	Supported (diminishing returns)
H2: Mediator (perceived effort)	–	Supported	Supported	Supported
H3: Moderator (product type)	Supported	Not supported	Not supported	–

FIGURES

Figure 1 – Study 1: Screenshot of Two App Reviews for “Fruit Ninja”

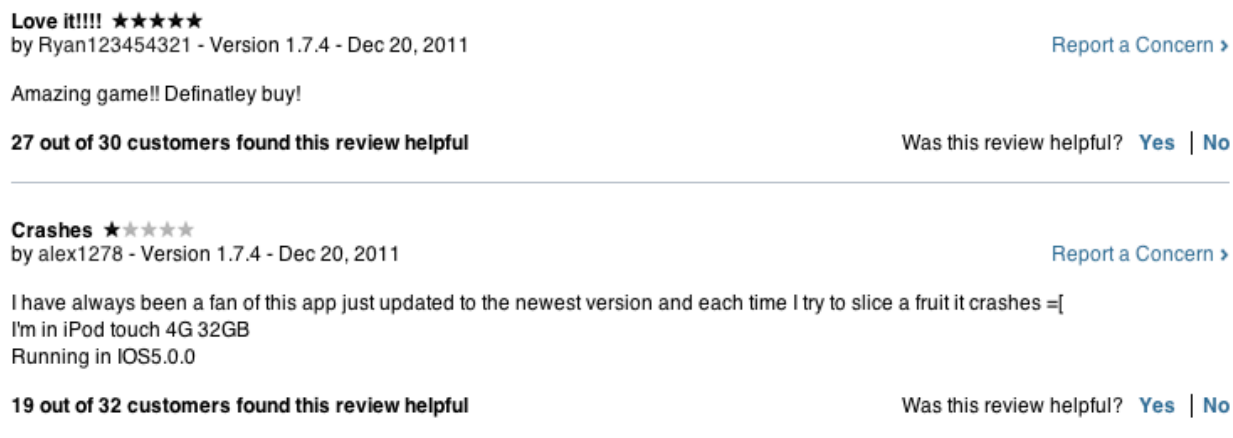
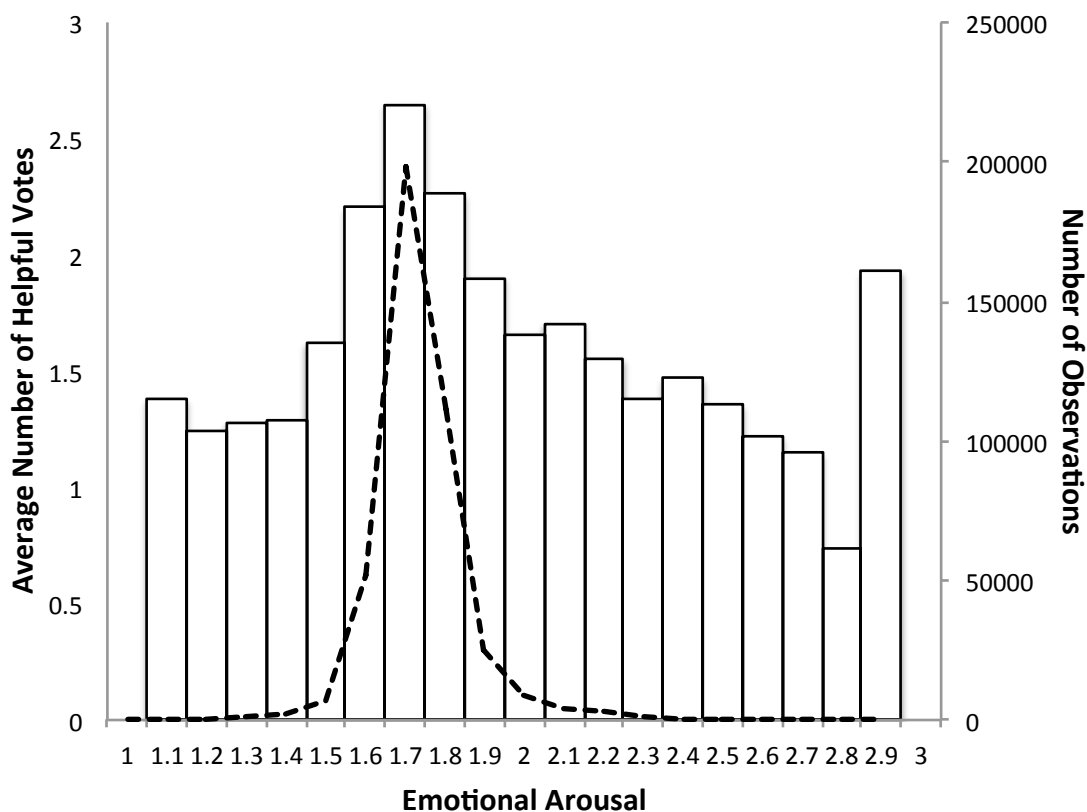
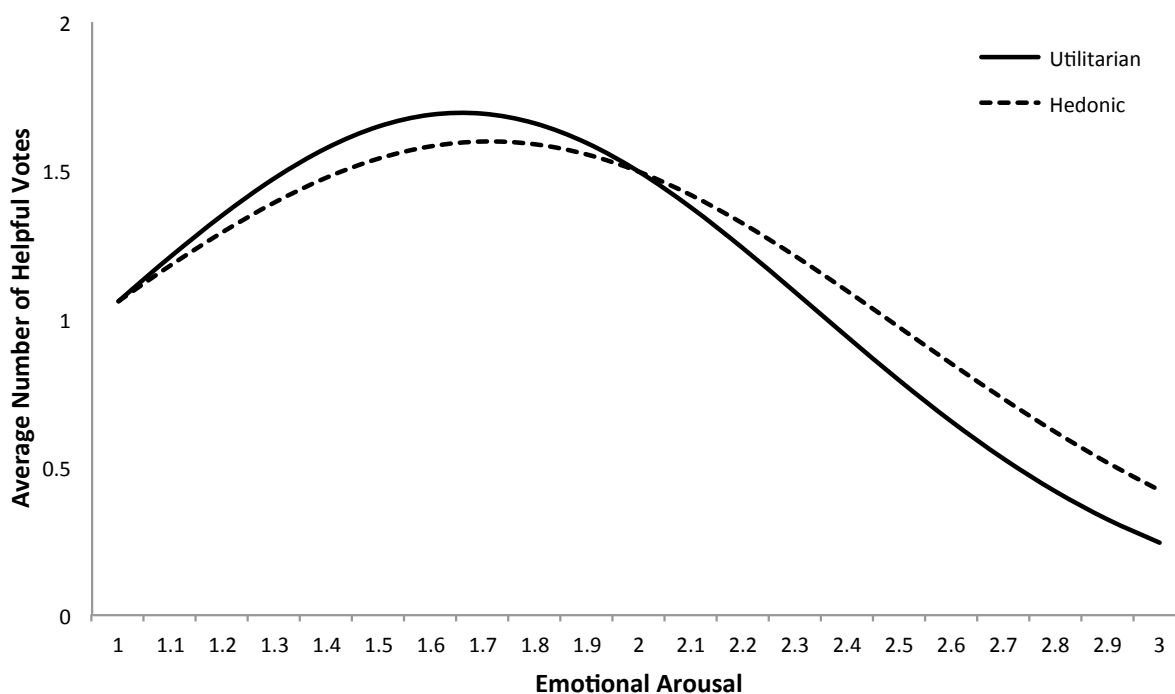


Figure 2 – Study 1: Review Helpfulness by Arousal



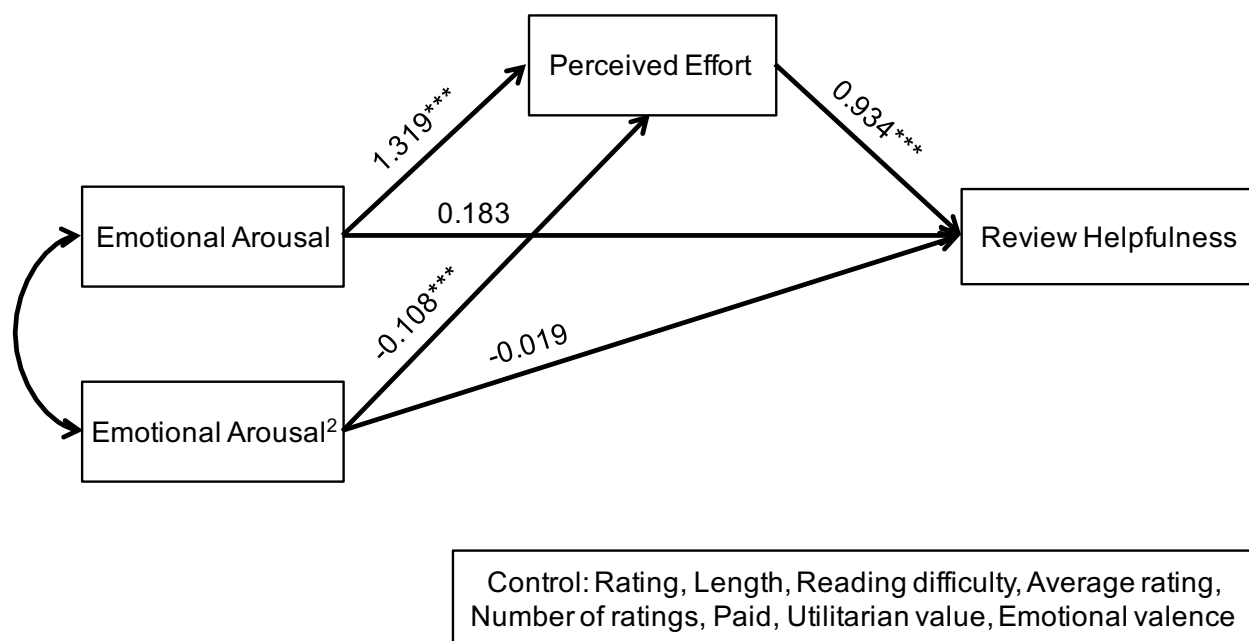
Notes: The height of each column represents the average number of helpful votes received by reviews at each 0.1 increment on the arousal measure. The dotted line indicates the number of observations for each increment.

Figure 3 – Study 1: Post-Hoc Plot of Moderation by Product Type



Notes: The curves depict expected values for review helpfulness based on Model 3 of Table 3. The utilitarian curve was generated using a utilitarian value one standard deviation above the mean (Utilitarian value = 4.15), and the hedonic curve was generated using a utilitarian value one standard deviation below the mean (Utilitarian value = -5.92). Mean values were used for all other variables in the model.

Figure 4 – Study 2: Nonlinear Mediation Model



Notes: *** significant at 0.01, ** significant at 0.05, * significant at 0.1

Figure 5 – Study 3: Perceived Helpfulness of Reviews by Arousal

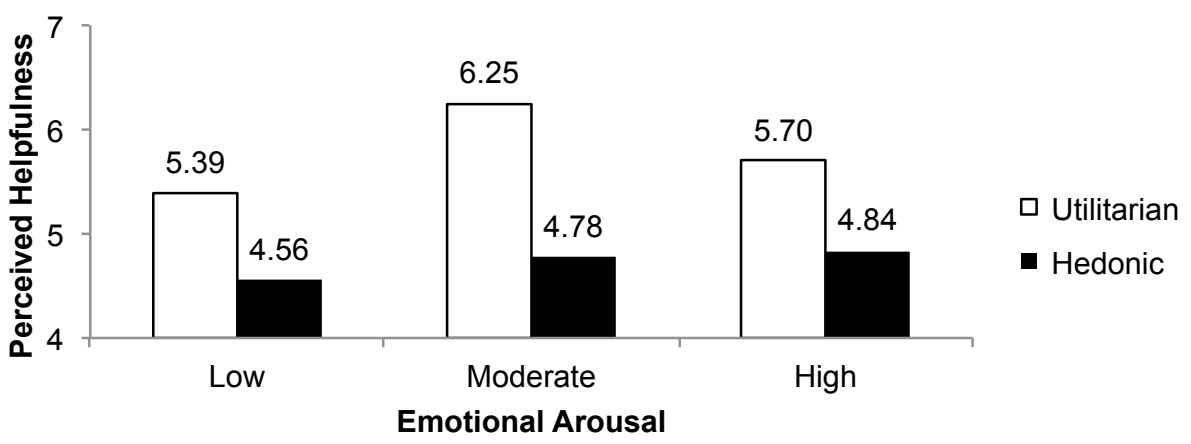
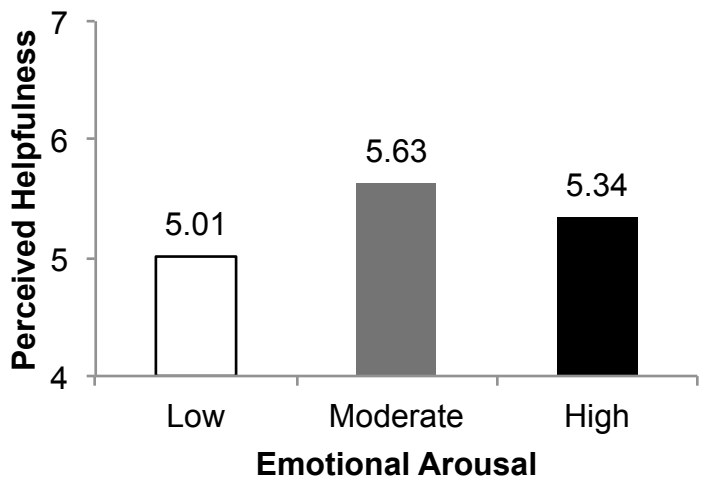


Figure 6 – Study 4: Perceived Helpfulness of Reviews by Arousal

