

When Consistency Matters: The Effect of Valence Consistency on Review Helpfulness

Simon Quaschnig

Department of Marketing, Ghent University, Tweeckerkenstraat 2, 9000 Ghent, Belgium

Mario Pandelaere

Department of Marketing, Ghent University, Tweeckerkenstraat 2, 9000 Ghent, Belgium

Iris Vermeir

Department of Marketing, Ghent University, Tweeckerkenstraat 2, 9000 Ghent, Belgium

When evaluating the helpfulness of online reviews, review valence is a particularly relevant factor. This research argues that the influence of review valence is highly dependent on its consistency with the valence of other available reviews. Using both field and experimental data, this paper shows that consistent reviews are perceived as more helpful than inconsistent reviews, independent of them being positive or negative. Experiments show that this valence consistency effect is driven by causal attributions, such that consistent reviews are more likely to be attributed to the actual product experience, while inconsistent reviews are more likely to be attributed to some reviewer idiosyncrasy. Supporting the attribution theory framework, reviewer expertise moderates the effect of consumers' causal attributions on review helpfulness.

Keywords: Customer Reviews, Review Valence, Consistency, Causal Attributions, Review Helpfulness, Reviewer Expertise.

doi:10.1111/jcc4.12106

Online reviews have become one of the most popular information sources for consumers (Chevalier & Mayzlin, 2006). However, not all reviews are perceived as equally helpful. Numerous factors can affect how helpful online reviews are perceived to be, including review and consumer aspects, as well as the characteristics of the reviewed product itself (e.g., Doh & Hwang, 2009; Mudambi & Schuff, 2010; Sen & Lerman, 2007; Willemsen, Neijens, Bronner, & de Ridder, 2011). One factor that is particularly relevant is the valence of the review (positive vs. negative). While the psychology literature consistently argues that negative information is more impactful (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001), research on online reviews has often demonstrated a positivity bias, where positive reviews are more helpful (Carlson & Guha, 2010; Pan & Zhang, 2011). The current research aims to understand the cause of this positivity bias, by investigating a factor that has been mostly ignored by

Editorial Record: First manuscript received on November 07, 2013. Revisions received on April 04, 2014 and July 29, 2014. Accepted by S. Shyam Sundar on September 13, 2014. Final manuscript received on September 19, 2014. First published online on November 10, 2014.

existing research: reviews are often consulted together with other reviews (see Cheung, Luo, Sia, & Chen, 2009; Cheung, Sia, & Kua, 2012; Purnawirawan, De Pelsmacker, & Dens, 2012 for an exception). Does the consistency of a review's valence with that of other available reviews affect the review's helpfulness (i.e., a valence consistency effect)?

We build on prior survey-based research that suggests a consistency effect may exist (Cheung et al., 2009; Cheung et al., 2012). *First*, by using data from Amazon (Study 1) and from two experiments (Study 2 and 3) we eliminate the potential influence of memory biases and unmeasured review characteristics on the effects in the existing survey-based studies (Cheung et al., 2009; Cheung et al., 2012). *Second*, the present research adds to the previous literature by illustrating the psychological processes underlying this valence consistency effect: Consistent reviews are attributed to the product's quality (Study 2 and 3), while inconsistent reviews are attributed to personal feelings of the reviewer. *Third*, we further test the role of causal attributions by investigating a factor that moderates the effect of causal attributions on review helpfulness: reviewer expertise. Review websites often identify reviewers as experts when they demonstrated their ability to write credible reviews in the past (Willemsen, Neijens, & Bronner, 2012). Due to perceptions of greater expertise of these reviewers, their reviews are seen as helpful, even when they are inconsistent with other available reviews; reviewer expertise thus attenuates the valence consistency effect (Study 3).

Finally, we discuss the role of our results in the broader context of the existing literature and illustrate the managerial contributions. Given the impact of reviews on websites like Amazon or Barnes & Nobles, it is important for companies to understand how review environments can be structured to improve the helpfulness of positive reviews and diminish the helpfulness of negative reviews.

Review Valence and the Consistency Effect

People often rely on others' opinions when making decisions. While the informational social influence theory (and the related conformity principle, the bandwagon effect and the social validity theory) suggests that information in online reviews is often expected to reflect the quality of a product (Burnkraut & Cousineau, 1975; Deutsch & Gerard, 1955), the perceived value of information for consumers' decisions (i.e., the review helpfulness; Mudambi & Schuff, 2010) determines the extent to which consumers will use this information. Dual-process theory has argued that both informational and normative factors can influence the helpfulness of online reviews (Cheung et al., 2009). *Informational factors* are related to the informational components of a review, including the content, source, and receiver of the information. Previous studies have focused on the length of the review, the credibility of the reviewer or the reviewed product as informational factors that can influence consumers' perceived review helpfulness (e.g., Doh & Hwang, 2009; Mudambi & Schuff, 2010; Sen & Lerman, 2007; Willemsen et al., 2011). Another particularly important informational factor is the review valence (i.e., whether the review is positive or negative; Cheung et al., 2009). *Normative factors* refer to the influences of others that are available in the review environment. As such, the review helpfulness is not based on the review itself, but on information of others' opinions (Cheung et al. 2009). For online reviews, a potentially relevant normative factor is the consistency between reviews, as it can moderate the effect of review valence.

Past research on information valence has observed a negativity bias, implying that negative information is often perceived as more valuable and has more impact on decisions relative to positive information (Baumeister et al., 2001). Based on the frequency-as-information perspective, negative information is often less common and therefore more attention-grabbing and more informative (Fiske, 1980; Peeters & Czapinski, 1990). Negative information is also more consequential (Rozin & Royzman, 2001) where in general, failing to act upon negative information (when it is correct) leads to worse outcome than failing to follow positive information (when it is correct; Baumeister et al., 2001). As such, purchasing a bad

product evokes more negative feelings than failing to purchase a good product. Moreover, the category diagnosticity theory (Skowronski & Carlston, 1987) argues that negative information primarily characterizes low quality products, while positive information can both characterize low- and high-quality products, which makes negative information more diagnostic for differentiating bad from good products. Surprisingly, research on the impact of online reviews has reported mixed findings (Chen & Lurie, 2013; Chevalier & Mayzlin, 2006; Yang & Mai, 2010) and studies on the *helpfulness* of Amazon reviews even show a positivity effect (i.e., positive reviews are perceived as more helpful than negative reviews; Carlson & Guha, 2010; Pan & Zhang, 2011). These studies draw on the confirmation bias theory (Russo, Meloy, & Medvec, 1998), which argues that consumers discredit negative information because it does not support their existing positive attitude to the reviewed product.

Existing research on review valence, however, has almost solely focused on the helpfulness of single online reviews, ignoring that websites, such as Amazon or Yelp, often present multiple reviews simultaneously. Hence, for most products, consumers are likely to find a mix of both positive and negative reviews online. As an exception, some studies showed that purely positive or negative *sets of reviews* are perceived as more helpful (Purnawirawan et al., 2012) and as more impactful on consumers' attitudes towards the product and website (Doh & Hwang, 2009; Lee, Park, & Han, 2008) than a *set* containing both positive and negative reviews. However, consumers usually vote on helpfulness of individual reviews rather than sets of reviews. In the current paper, we therefore argue that when consumers consult information from multiple reviews, the consistency of a review's valence with the valence of other available reviews will determine its helpfulness.

The Role of Attributions

The effect of review valence consistency can operate in one of two directions. First, according to the frequency-as-information principle, infrequent information is often more diagnostic. Hence, a review that is inconsistent with other available reviews will be perceived as more helpful than a review that is consistent with other reviews. An inconsistent review will be perceived as more informative and diagnostic (Skowronski & Carlston, 1987), which could translate into higher review helpfulness (Jiang & Benbasat, 2007; Wu, 2013). This reasoning would predict a negativity bias in the overly positive environment of websites such as Amazon.com. As mentioned before, however, studies almost consistently reported a positivity effect when looking at real online reviews (e.g., Carlson & Guha, 2010; Pan & Zhang, 2011).

Alternatively, attributional processes may cause a review that is consistent with most of the other available reviews to be perceived as more helpful than an inconsistent review. When facing information, such as online reviews, people make inferences about the validity of the expressed opinions (Folkes, 1988; Mizerski, 1982). Essentially, reviews can be either attributed to the actual product experience or to reviewer-related motivations, traits, or attitudes (Chen & Lurie, 2013; Monga & John, 2008; Sen & Lerman, 2007). Consumers' causal attributions then influence the perceived helpfulness of online reviews. The more a review is attributed to a product's factual performance or quality, the more a consumer will evaluate the review as legitimate, believable, and, hence, as helpful (Mizerski, 1982). In contrast, when consumers attribute the review to some reviewer idiosyncrasy, the review will be perceived as less helpful (Chen & Lurie, 2013; Sen & Lerman, 2007).

A set of online reviews conveys information about the consensus between reviewers about the consumption experience and, hence, the consistency between the reviews (Laczniak, DeCarlo, & Ramaswami, 2001). In a case of high consistency in valence between reviews, a review is more likely to be attributed to the product experience, while inconsistent reviews are more likely attributed to the

personal motivations of the reviewer (Fiske & Taylor, 1991; Kelley, 1967; Kelley, 1973). Consequently, we hypothesize that review valence consistency leads to higher product attributions causing the review to be perceived as more informative and, hence, as more helpful. Therefore, we hypothesize:

H1. When a review's valence is consistent (vs. inconsistent) with the valence of other available reviews, the review will be perceived as more (vs. less) helpful.

H2. Causal attributions mediate the effect of valence consistency on review helpfulness. Stronger valence consistency causes more product-related and less reviewer-related attributions, which increases review helpfulness.

We further test the attribution theory explanation by examining how reviewer expertise moderates the valence consistency effect. Many review websites publicly label reviewers that have a record of writing credible reviews as experts. This so-called *rated expertise* has the potential of affecting people's perception about the believability of a review (Willemsen, Neijens, & Bronner, 2012). Expert reviews are often perceived as more trustworthy and informative, as they are perceived as a more concrete representation of reality (Schlosser, 2011; Willemsen et al., 2012). As such, rated expertise signals a higher validity of review information (Eastin, 2001) and reduces readers' suspicion about the possibility that a review is driven by the reviewer's personal motivations (Willemsen et al., 2012). Consequently, rated expertise is a strong cue that might override the effect of a review's valence consistency and related causal attributions and, hence, the valence consistency effect. Given the higher trustworthiness of expert reviews, the review might be seen as very helpful, even when it is inconsistent with other reviews.

H3. Rated reviewer expertise moderates the extent to which valence consistency affects review helpfulness. Inconsistent expert reviews will be perceived as more helpful than regular inconsistent reviews.

Three studies test our theoretical framework. Study 1 uses real life data from Amazon to provide evidence for the existence of the valence consistency effect. Next, an experiment illustrates that consistent reviews are attributed to product-related factors, while inconsistent reviews are attributed to non-product-related factors, which in turn explains the difference in helpfulness (Study 2). Study 3 further investigates the underlying mechanism by illustrating that the valence consistency effect disappears for reviews written by experts.

Study 1: The Valence Consistency Effect on Amazon

Study 1 investigates the influence of valence consistency on review helpfulness. Additionally, this study provides first evidence in favor of the attribution theory and against the frequency-as-information account. While we expect that valence consistency would cause a higher helpfulness, the frequency-as-information theory suggests that inconsistent information is more diagnostic, which would result in a negative effect. Using a quasi-experimental design, we look at the helpfulness of book reviews gathered from Amazon.com. We chose this data source, since Amazon is one of the largest online retailers with one of the most active reviewing communities online. Their bidirectional WOM network not only allows consumers to provide and read reviews, but also to judge the helpfulness of the review message. Also, reviews are presented simultaneously, with multiple reviews being available on the same screen.

Method

Review data of 1,300 online reviews for 117 different books was extracted from Amazon.com (with a maximum of 20 random reviews per product). To ensure that both relatively popular and unpopular

books are taken into account, we included books from different sales ranks, with group of products ranked 1st to 25th, 101st to 125th, 501st to 525th, 5,001st to 5,025th and 10,001st to 10,025th.

The dependent variable in our model is the perceived review helpfulness, measured by the proportion of readers that found a review to be helpful. This variable was operationalized by dividing the number of people who evaluated the review as helpful by the total votes in response to the “was this review helpful to you” question on Amazon.com. As main independent variables of interest, we extracted the star score given in the reviews as the *review valence* and the average product score as a proxy for the average star score of other available reviews (both on a scale of 1 to 5; with 5 being a very positive evaluation). We used the average product score as a proxy, since the composition of the displayed reviews is often changing and it is not clear which reviews were displayed when consumers evaluated the reviews’ helpfulness. Consistent with previous research (e.g., Chevalier & Mayzlin, 2006), the average online customer review in our data set is positive ($M = 4.21$, $SD = 1.23$), as is the average product score ($M = 4.12$, $SD = .60$). The review length and the sales rank were included as control variables. *Review length* was operationalized as the number of words a review contains ($M = 141.18$, $SD = 108.98$). *Sales rank*, measured at the moment of data collection, ranges from 1 for the top-sold product to 10,025 ($M = 2141.99$, $SD = 124.00$). Since both variables are positively skewed (a skewness of 1.21 for review length and 1.49 for sales rank), they were log transformed (base 10 log transformation; Dallal, 1999).

Including multiple reviews for each book means that reviews are nested within books and review length and review valence vary within books, rather than merely across books. Therefore, we analyzed the data using a multilevel regression analysis, requiring the specification of an appropriate error structure. The most suitable error covariance structure was compound symmetrical, implying constant correlation between every pair of errors. The estimation of the error degrees of freedom in the statistical tests relied on Satterthwaite’s approximation, which may result in fractional degrees of freedom (Littell, Stroup, & Freund, 2002). The interpretation of the parameter estimates is the same as with ordinary linear regression; merely the standard errors of the parameters are adjusted to obtain correct test statistics. The regression model also included interactions with the control variables review length and sales rank, resulting in two three-way interactions (*review valence x average product score x review length* and *review valence x average product score x sales rank*). To facilitate the interpretation of the interactions, we standardized the average product score, the review length and the sales rank (Dawson, 2014).

Results

The results (see Table 1) show that on average, consumers evaluate positive reviews as more helpful than negative reviews (as indicated by the main effect of review valence). In addition, a significant two-way interaction between the review score and the average product score shows that the helpfulness of a review is also affected by the consistency of its valence with the valence of other available reviews. A simple slope analysis shows that when the average score is positive (1 SD above the mean; a score of 4.72), positive reviews are seen as more helpful than negative reviews ($\beta = 13.82$, $t(1277.54) = 12.66$, $p < .001$). In contrast, when the average score is moderate (1SD below the mean; a score of 3.52), indicating a fair share of both positive and negative reviews, the valence of a review no longer has an effect on the perceived review helpfulness ($\beta = -.50$, $t(1277.15) = -.62$, $p = .54$). Since the number of negatively evaluated products on Amazon is very low, we are unable to look at the results for a negative context.

Review length does not affect the valence consistency effect. While on average, a longer review is perceived as more helpful (as indicated by the main effect) we find no significant two-way interaction with either review valence or the average product score and no significant three-way interaction with review valence and the average score. In contrast, we find that the valence consistency effect is affected by the sales rank of the reviewed product. A significant three-way interaction indicates that the valence consistency effect is less strong for better-selling books. When the sales rank is relatively low (1 SD

Table 1 Multilevel Regression Results for the Prediction of Review Helpfulness (Study 1)

Parameter	Estimate	Standard Error	t-value	p-value
Intercept	36.44	3.48	10.48	< .001
Review valence	6.67	.73	9.18	< .001
Average score	-24.31	2.84	-8.55	< .001
Review valence x average score	7.18	.63	11.37	< .001
<i>Control variables</i>				
Review length	11.99	3.24	3.70	<.001
Review length x review valence	-.58	.72	-.80	.43
Review length x average score	2.73	2.52	1.08	.28
Review length x review valence x average score	-.25	.61	-.40	.69
Sales rank	3.13	3.59	.87	.38
Sales rank x review valence	.60	.75	.80	.42
Sales rank x average score	3.21	2.79	1.15	.25
Sales rank x review valence x average score	-1.36	.63	-2.17	.030

Note: Average score is standardized; Sales rank and review length are log transformed and subsequently standardized

below the mean, indicating a better selling product), the interaction between review valence and average score is weaker ($\beta = 6.07$, $t(1255.07) = 6.11$, $p < .001$) than when the sales rank is relatively high (1 SD above the mean; $\beta = 7.28$, $t(1266.55) = 6.63$, $p < .001$). In either case, however, the valence consistency effect prevails.

Discussion

The results of Study 1 provide initial evidence for the valence consistency effect (H1). The Amazon data reveal that the average product score, as a proxy of the valence of other available reviews, influences the relationship between the valence of the focal review and the review helpfulness. As such, a positive review for an on average positively evaluated product is perceived as more helpful than a negative review would be. This higher helpfulness for positive reviews disappears when there is an equal share of positive and negative reviews. The positive effect of the valence consistency on review helpfulness also suggests that the frequency-as-information explanation is unlikely.

Drawing on 1,300 reviews from over 100 different books, the first study is interesting because of its external validity. There are, however, some limitations to this study. First, since the average online customer review and the average product score are positive, we are unable to draw conclusions about situations where most of the available reviews are negative. Second, the average score that we used as a proxy for the valence of other presented reviews might not be an exact representation of the reviews the consumers actually looked at when evaluating the focal review. Moreover, only a small portion of people who read the reviews also vote on the review helpfulness and there might be other unobservable and uncontrollable variables related to reviewers, readers and the review content, which could influence the review helpfulness (Wu, 2013). To address these issues and to examine the underlying mechanism of the valence consistency effect, we conduct two experiments.

Study 2: An Attribution Theory Explanation

Study 2 aims to replicate the findings of the Amazon study with an experimental design and to test the proposed process underlying the effect of valence consistency on review helpfulness. As proposed

in H2, we expect that inconsistent reviews are attributed to the reviewer and that consistent reviews are attributed to the product experience. Since product-related factors are more informative about the actual quality of a product, consistent reviews should be perceived as more helpful.

Method

Hundred and sixty undergraduate students from Ghent University (mean age = 20.32 years, $SD = 3.04$; 132 women) were randomly assigned to the conditions of a 2 (review valence) \times 3 (context valence) between-subjects design. As stimuli, we presented the respondents a screenshot of a review website presenting three filler reviews and one focal review. Every respondent received the same four reviews, in either positive or negative versions. Negative reviews were created by replacing the positive adjectives in the positive reviews with negative ones. The focal review was either negative (one star review) or positive (five star review). Context valence was manipulated through the valence of the filler reviews. This resulted into a positive, neutral and negative context. Including the (either positive or negative) focal review, the *positive context* consisted of three positive and one negative reviews, the *neutral context* consisted of two positive and two negative reviews and the *negative context* consisted of three negative and one positive reviews (see Appendix 1 for an example). While an alternative option of using all positive and all negative reviews to create consistent conditions was viable, we prefer our current choice as it creates consistent and inconsistent conditions that are equidistant from the neutral condition. In both cases, the difference is created by changing the valence of a single review. The alternative option would have led to a situation where the inconsistent condition would differ from the neutral condition by a single review, but from the consistent condition by two reviews.

A pretest ($n = 60$; mean age = 35.72 years, $SD = 12.85$; 26 women) indicated on a 7-point scale that the focal negative review is significantly more negative than the focal positive review ($M_{neg} = 6.00$ vs. $M_{pos} = 1.90$, $F(1, 58) = 123.21$, $p < .001$). Moreover, the context valence (Goldsmith, Lafferty and Newell, 2000; I have the impression that most people: are satisfied/dissatisfied with the restaurant, find the restaurant good/bad, have a positive/negative opinion about the restaurant, would/would not recommend the restaurant; $\alpha = .966$), was perceived as more negative in the negative context and more positive in the positive context ($M_{pos} = 2.38 < M_{neut} = 4.20 < M_{neg} = 6.14$, $F(2, 57) = 123.21$, $p < .001$). A Bonferroni post hoc test indicates significant differences between each group (all p 's $< .001$).

After reading the reviews, the participants evaluated the focal review's helpfulness on a 7-point scale, with 1 being "not helpful at all" and 7 being "very helpful" ($M = 4.81$, $SD = 1.19$). After that, we assessed product- and reviewer attributions using measures adapted from Laczniak et al. (2001). Product attributions were measured by asking participants how much they agree with five statements ("The restaurant is an inferior/a superior restaurant," "The restaurant is unpopular/popular," "The restaurant performed poor/good," "The restaurant is unusual/usual," and "The restaurant lacked/had the necessary qualities"; Cronbach's $\alpha = .85$), and reviewer attributions were assessed with four statements ("He doesn't/does know enough about restaurants," "He doesn't/does appear to have the expertise to evaluate the restaurant properly," "He wanted/didn't want to look smarter than he really is," and "He is the type of person who always says good/bad things"; Cronbach's $\alpha = .79$). The attributions were measured on a 7-point scale, with 1 being "do not agree at all" and 7 being "agree completely." Subtracting reviewer attributions from product attributions resulted in the causal attribution score ($M = -.29$, $SD = 1.67$). A positive (negative) score indicates that the product (reviewer) attributions outweigh the reviewer (product) attributions. Inclusion of gender as control variable (covariate) did not alter the effects on review helpfulness or causal attributions.

Results

Review Helpfulness.

The results show a significant two-way interaction between the valence of the focal review and the average valence of the context ($F(2,154) = 12.43, p < .001$; see Figure 1a). The *positive review* is significantly more helpful in a positive context than in a negative context ($M_{\text{pos}} = 5.46$ vs. $M_{\text{neg}} = 4.42, F(1, 154) = 11.36, p < .001$), marginally significantly more helpful in a positive context than in a neutral context ($M_{\text{pos}} = 5.46$ vs. $M_{\text{neut}} = 4.93, F(1, 154) = 3.08, p = .081$) and marginally significantly more helpful in a neutral context than in a negative context ($M_{\text{neut}} = 4.93$ vs. $M_{\text{neg}} = 4.42, F(1, 154) = 2.71, p = .10$). The *negative review* is more helpful in a negative context than in a positive context ($M_{\text{neg}} = 5.12$ vs. $M_{\text{pos}} = 4.09, F(1, 154) = 11.98, p < .001$) and more helpful in a neutral context than in a positive context ($M_{\text{neut}} = 5.00$ vs. $M_{\text{pos}} = 4.09, F(1, 154) = 9.13, p = .003$). No significant difference, however, was found between the negative context and the neutral context ($M_{\text{neg}} = 5.12$ vs. $M_{\text{neut}} = 5.00, F(1, 154) = .14, p = .71$).

The Role of Causal Attributions.

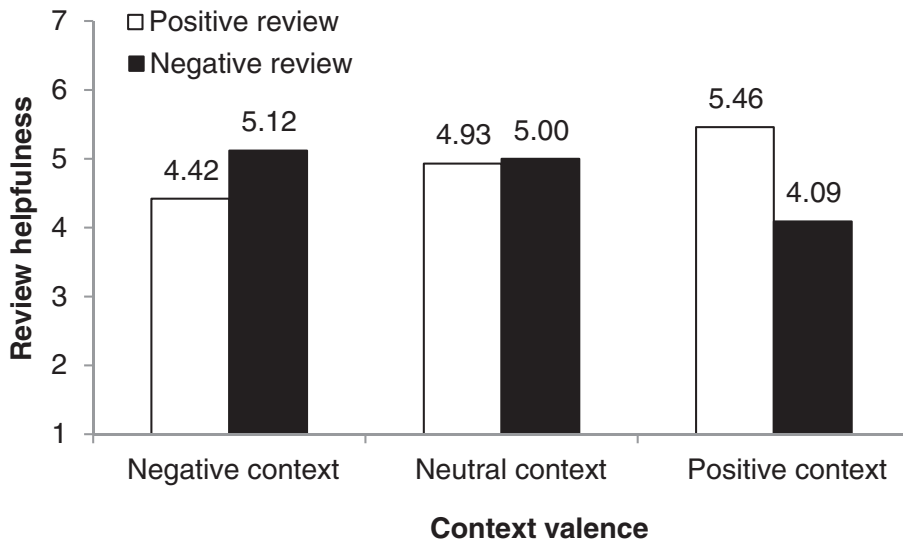
We find a significant two-way interaction between the valence of the focal review and the average valence of the context ($F(2,154) = 42.70, p < .001$; see Figure 1b). The *positive review* is significantly more attributed to the product experience when most other reviews were positive, while it is more attributed to the reviewer when most other reviews were negative ($M_{\text{pos}} = .88$ vs. $M_{\text{neg}} = -1.08, F(1, 154) = 29.84, p < .001$). Moreover, we find a significant difference between the neutral and the negative condition ($M_{\text{neut}} = .29$ vs. $M_{\text{neg}} = -1.08, F(1, 154) = 14.91, p < .001$), but only a marginally significant difference between the positive and the neutral context valence ($M_{\text{pos}} = .88$ vs. $M_{\text{neut}} = .29, F(1, 154) = 2.73, p = .10$). Similarly, *negative reviews* cause more product attributions when the context is negative rather than positive ($M_{\text{neg}} = .50$ vs. $M_{\text{pos}} = -2.07, F(1, 154) = 55.23, p < .001$). Also, there is a significant difference between the neutral context and the positive context ($M_{\text{neut}} = .20$ vs. $M_{\text{pos}} = -2.07, F(1, 154) = 42.48, p < .001$), while no significant difference was found between the negative and neutral context ($M_{\text{neg}} = .50$ vs. $M_{\text{neut}} = .20, F(1, 154) = .61, p = .44$).

Supporting H2, respondents' causal attributions explain the valence consistency effect. We coded *valence consistency* as low (i.e., a positive review in a negative context or a negative review in a positive context), moderate (i.e., a positive or negative review in a neutral context), and high (i.e., a positive review in a positive context or a negative review in a negative context). As reviews are more consistent with other reviews, they are perceived as more helpful ($\beta = .43, t(156) = 2.64, p = .009$), which was unaffected by review valence ($\beta = .09, t(156) = .39, p < .70$). Similarly, valence consistency has a positive effect on the causal attributions ($\beta = 1.16, t(156) = 5.68, p < .001$), unaffected by review valence ($\beta = -.18, t(156) = -.63, p = .53$). Causal attributions also significantly predict the perceived helpfulness of the focal review ($\beta = .43, t(158) = 9.50, p < .001$). In a simultaneous regression predicting the review helpfulness, the causal attributions continue to be a significant predictor ($\beta = .42, t(157) = 8.07, p < .001$), whereas review consistency is not any more ($\beta = .02, t(157) = .22, p = .83$). The bootstrap analysis for the indirect effect shows that this mediation is significant (95% CI [.32, .61]).

Discussion

Consistent with the findings of our Amazon study (Study 1), the results show that a positive review is evaluated as more helpful in a positive context than in a negative context. Conversely, a negative review is evaluated as more helpful in a negative context than in a positive context. In a neutral context, we see that the effect of review valence disappears. Moreover, we find evidence for H2, by providing support for

(a) Effect on helpfulness



(b) Effect on causal attributions

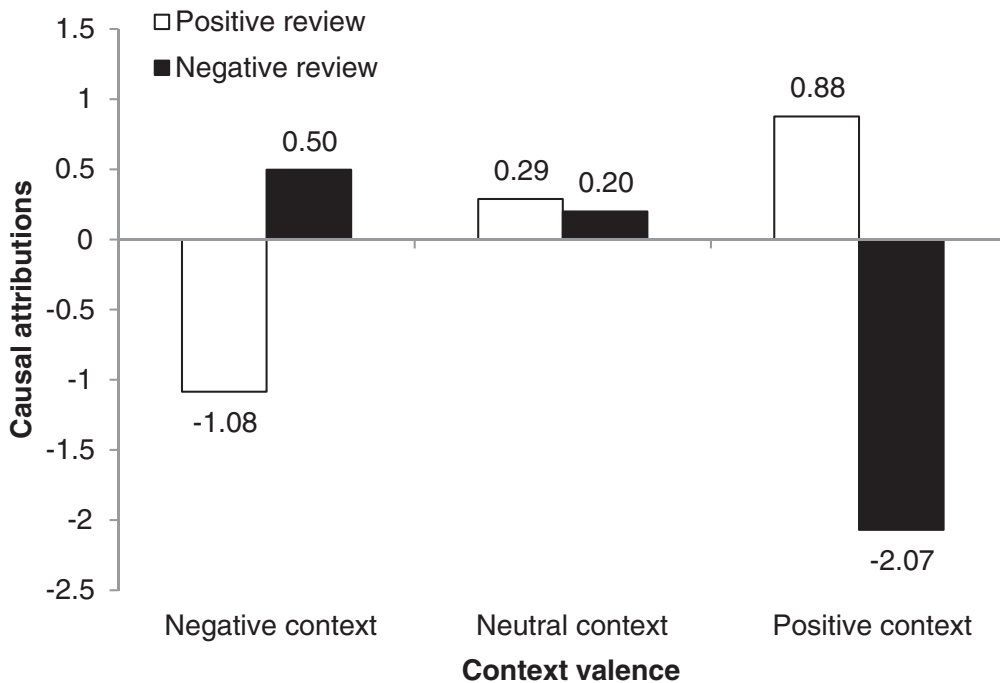


Figure 1 Effect of Review Valence and Context Valence on Review Helpfulness and Causal Attributions (Study 2)

the attribution theory explanation. Participants attribute the review in a consistent context to the actual product performance while the reviews in an inconsistent context are attributed to personal factors of the reviewer. These causal attributions, in turn, mediate the valence consistency effect.

Study 3: The Role of Reviewer Expertise

Study 3 further tests the attribution theory hypothesis by investigating the moderating role of reviewer expertise (H3). Other than on valence consistency, people also rely on cues about the reviewer expertise. Experts are often perceived as more objective than other reviewers (Willemsen et al., 2012). Consequently, we expect that rated expertise overrides the effect of people's causal attributions on review helpfulness. Hence, an inconsistent expert review might still be perceived as highly helpful, making the valence consistency effect disappear.

Method

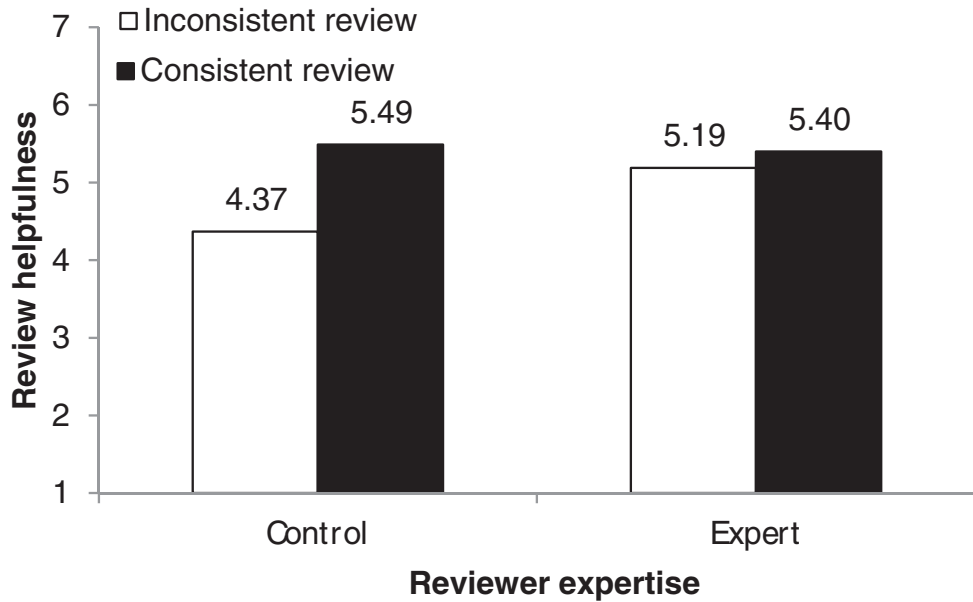
One hundred and forty people from an online panel consisting of both students and nonstudents (mean age = 22.07 years, $SD = 6.77$; 93 women) participated in a 2 (review consistency) \times 2 (reviewer expertise) between-subjects design. Respondents were given a screenshot of four reviews, consisting of three filler reviews and one focal review. Since no significant differences were found in the valence consistency effect for the positive and the negative reviews (see Study 2), we only use a positive (five-star review) focal review. Review consistency was manipulated through the valence of the filler reviews. The filler reviews were either two positive reviews and one negative review (the consistent condition) or three negative reviews (the inconsistent condition). The negative filler reviews were created by replacing the positive adjectives in the positive reviews with negative ones. As in Study 2, a pretest ($n = 35$; mean age = 25.49 years, $SD = 8.55$; 25 women) indicated that the inconsistent condition was more negative, and hence inconsistent, than the consistent condition ($M_{\text{cons}} = 1.70$ vs. $M_{\text{incons}} = 6.00$, $F(1,33) = 150.01$, $p < .001$). Similar to previous research (Willemsen et al., 2012), reviewer expertise was manipulated by indicating the focal review as 'expert review'. Additionally, the reviewer identified himself as an IT consultant.

Participants were asked to indicate their perceived helpfulness of the focal review on a 7-point scale, with 1 being "not helpful at all" and 7 being "very helpful" ($M = 5.09$, $SD = 1.26$). We measured the respondents' causal attributions with the same scale as in Study 2 (Lacznik et al., 2001), adapted to control for the conceptual overlap with the expert manipulation.¹ We calculated a causal attribution score ($M = .28$, $SD = 1.67$) by subtracting reviewer attributions (Cronbach's $\alpha = .64$) from product attributions (Cronbach's $\alpha = .88$). Controlling for the effect of gender on the dependent variables of helpfulness and causal attributions did not change the results.

Results

Review Helpfulness. As illustrated in Figure 2a, the results indicate a significant two-way interaction between review consistency and reviewer expertise ($F(1,136) = 5.03$, $p = .026$). Replicating the consistency effect, the *regular* focal review is evaluated as less helpful in the inconsistent condition than in the consistent review. ($M_{\text{cons}} = 5.49$ vs. $M_{\text{incons}} = 4.37$, $F(1,71) = 13.50$, $p < .001$). For the *expert review*, however, the perceived helpfulness is similarly high in the consistent and the inconsistent condition ($M_{\text{cons}} = 5.40$ vs. $M_{\text{incons}} = 5.19$, $F(1,65) = .67$, $p = .42$). Other contrasts illustrate that when inconsistent with other available reviews, the expert review is significantly more helpful than the regular review ($M_{\text{expert}} = 5.19$ vs. $M_{\text{regular}} = 4.37$, $F(1, 136) = 8.22$, $p = .005$). In the consistent condition, both reviews are equally helpful ($M_{\text{expert}} = 5.40$ vs. $M_{\text{regular}} = 5.49$, $F(1, 136) = .09$, $p = .76$).

(a) Effect on helpfulness



(b) Effect on causal attributions

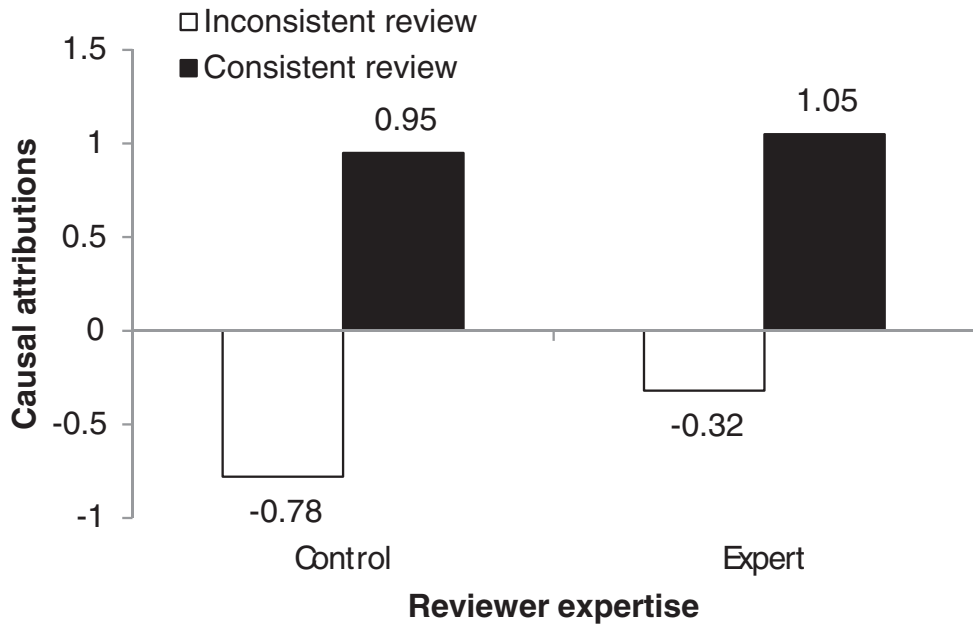


Figure 2 Effect of Review Consistency and Reviewer Expertise on Review Helpfulness and Causal Attributions (Study 3)

The Role of Causal Attributions. As illustrated in Figure 2b, valence consistency of the *regular review* increases the causal attributions score, as product attributions increase and reviewer attributions decrease ($M_{cons} = .95$ vs. $M_{incons} = -.78$, $F(1,71) = 20.66$, $p < .001$). Similarly, for the *expert review*, valence consistency also increases the causal attribution score ($M_{cons} = 1.05$ vs. $M_{incons} = -.32$, $F(1,65) = 15.15$, $p < .001$).

A moderated mediation analysis (Model 15, Hayes, 2012) tested the impact of the expert review on the valence consistency. For the regular review, causal attributions, significantly predict the perceived helpfulness of the review ($\beta = .49$, $t(71) = 6.83$, $p < .001$). In a simultaneous regression prediction the review helpfulness of the regular review, the causal attributions continue to be a significant predictor ($\beta = .44$, $t(70) = 5.45$, $p < .001$), whereas review consistency is not ($\beta = .36$, $t(70) = 1.25$, $p = .22$). Conditional indirect effects indicate that the mediation is significant for regular reviews (95% CI [.34, 1.11]). For expert reviews, in contrast, causal attributions had no significant effect on review helpfulness ($\beta = .14$, $t(65) = 1.51$, $p = .14$). Moreover, we did not observe a consistency effect to start with and no significant mediation occurs (95% CI [-.10, .57]).

Discussion

The results of Study 3 further support our theoretical framework. Replicating the findings of Study 2, we show that causal attributions explain the valence consistency effect. Second, we support our third hypothesis by showing that the effect of valence consistency on review helpfulness is moderated by reviewer expertise. For regular reviews, inconsistency makes a review less helpful. Expert reviews, in contrast, are always evaluated highly helpful, regardless of valence consistency. Further analysis indicated that reviewer expertise attenuates the effect of causal attributions on perceived review helpfulness.

General Discussion

The valence of a review is a particularly important factor for the perceived helpfulness of online reviews. Previous research has suggested that positive online reviews are often perceived as more helpful, because they confirm consumers' existing preferences (Carlson & Guha, 2010; Pan & Zhang, 2011). However, existing research mostly ignored that consumers often do not read a single review in isolation. The current research argues that in a context with multiple reviews, the consistency of a review's valence with other available reviews, rather than a review's valence alone, will determine the perceived helpfulness. These findings also suggest people's causal attributions as an explanation for the valence bias.

Using both secondary data and experiments, we show that the consistency of a review's valence with other available reviews determines a review's helpfulness. An analysis of reviews from Amazon.com (Study 1) shows that positive reviews are on average more helpful. Taking the consistency with other reviews into account, however, shows that the positivity effect is driven by the overall positivity of Amazon reviews. Review valence is unrelated to helpfulness, when there is a fair share of both positive and negative reviews present. Additional studies (Studies 2 and 3) replicated the consistency effect in experimental settings and showed that both positive and negative reviews are more helpful when their valence is consistent with that of other reviews.

Subsequent experiments provide evidence for the attribution theory explanation (Studies 2 and 3). When a review's valence is consistent with other available reviews, the review is perceived to be reflective of a product's actual quality. Inconsistent reviews, in contrast, are attributed to reviewer factors, that are considered unrelated to a product's quality and, hence, irrelevant for the purchase decisions.

These causal attributions, in turn, affect review helpfulness. Since valence consistency results in higher review helpfulness, the current findings also suggest that a frequency-as-information explanation is unlikely.

Finally, the influence of valence consistency on review helpfulness is moderated by the reviewer's expertise (Study 3). In particular, we show that the valence consistency effect disappears for reviews written by expert reviewers. Expert cues override the effect of causal attributions on review helpfulness. Compared to a *regular* review, causal attributions no longer have a significant impact on the helpfulness of an expert review.

Contributions and Directions for Future Research

Valence biases are well documented. Often negative information has both more impact and is perceived as more helpful (negativity bias, e.g., Baumeister et al., 2001; Rozin & Royzman, 2001). However, research that investigates valence effects for online reviews on websites such as Amazon.com, demonstrates that positive reviews are often more helpful (positivity bias, Carlson & Guha, 2010; Pan & Zhang et al., 2010). The current paper extends prior survey-based research on valence consistency (Cheung et al., 2009; Cheung et al., 2012) with data from real reviews and experiments and demonstrates that valence consistency determines whether a positivity or negativity bias is obtained for online reviews.

Our findings offer an explanation for the positivity bias found for reviews on Amazon. Our results provide some evidence that negative information has more impact on people's helpfulness than positive information. A negative review is only evaluated as less helpful when the majority of other reviews are positive; while a positive review is already less helpful when the positive reviews are not in the majority. On many websites, however, the majority of online reviews are positive, which renders positive reviews as more consistent with the average review valence and, in turn, as more helpful. As such, the consistency bias offers an explanation of why many studies using Amazon-data find a positivity bias.

Moreover, by establishing the causal attribution theory as underlying reason of the valence consistency effect, our research deviates from previous research on valence effects. The positivity effect found for Amazon reviews is often explained with the confirmation bias theory (Russo, Meloy, & Medvec, 1998), arguing that consumers read reviews for products for which they have an existing preference and discredit negative information to support this preference (Carlson & Guha, 2010; Pan & Zhang, 2011). The current findings show, however, that even in the absence of an existing preference, the abundance of positive reviews on Amazon will cause a positivity bias, due to higher product attributions.

Our results also deviate from the frequency-as-information perspective, which argues that infrequent events are more attention-grabbing and more diagnostic (Fiske, 1980; Peeters & Czapinski, 1990). As such, our findings also deviate from the findings of Wu (2013), where inconsistent reviews were found to be more helpful. However, the research of Wu (2013) featured a temporal cue ("*the latest review*"). Possibly, such temporal cues may suggest the presence of a trend, or a recent change in product quality. As such, a recent negative review for an overall positive product might be seen as more helpful, since it is more diagnostic for the current quality. Future research should further explore whether temporal cues attenuate the valence consistency effect.

By offering an attribution theory explanation, the current findings also contribute to the emerging literature that illustrates the role of causal attributions for the helpfulness of individual reviews. For example, Sen and Lerman (2007) demonstrated that negative reviews for utilitarian products (i.e., products that provide functional benefits) are more readily attributed to product-related factors than negative

reviews for hedonic products (i.e., products that provide an affective and sensory experience and yield pleasure), resulting in a higher perceived helpfulness. Another research illustrated that positive reviews are more attributed to the reviewer (vs. the product) than negative reviews and that indicating that a review was written shortly after a product experience reduces reviewer-related attributions for positive reviews (Chen & Lurie, 2013). By establishing the causal attribution theory as underlying reason of the valence consistency effect, the current research broadens our understanding about the role of causal attributions in review environments.

Finally, the current paper contributes to the literature on reviewer expertise. Previous research has argued that expert reviews are seen as more trustworthy and helpful (e.g., Willemsen et al., 2012). Our findings illustrate that expert reviews are able to withstand the impact of inconsistency on the review's helpfulness: While inconsistent reviews always cause more reviewer-related attributions, reviewer expertise overrides the effect of causal attributions on review helpfulness. Consequently, an expert review is always considered helpful. The current results also imply that expert reviews are only more helpful than regular reviews, as long as they are inconsistent with other reviews, while for consistent reviews, reviewer expertise does not make a difference.

The current findings also have important managerial implications. Online reviews are not only a popular decision aid for consumers and next to independent review websites, consumers often read reviews on commercial websites such as Amazon.com or Barnes & Nobles (Chevalier & Maylin, 2006). These companies are often concerned about the excessive impact of negative reviews on their sales. While companies might not be able to control the reviews that are posted on their website, the current findings suggest that they can focus on the way reviews are presented to boost the impact of positive reviews and to reduce the impact of negative reviews. Our results suggest that the helpfulness of a positive review might be strongly affected by the presence of several negative reviews. In contrast, one might consider presenting negative reviews between positive reviews, in order to diminish the impact of the former. Our findings also stress the importance of positive expert reviews, as they are more resistant against negative reviews.

Finally, future research should further explore the mechanism by exploring possible boundary conditions for the consistency effect. The salient availability of other reviews may facilitate a focus on the set of reviews as a whole (i.e., holistic processing). Hence, a review's consistency with other reviews becomes important. In contrast, when people are induced to approach a review analytically and thus focus on the individual components of a particular review, they might judge the review more readily on its inherent, content-related qualities. The same review might then be evaluated more independently of its consistency with other reviews, attenuating the valence consistency effect. Future studies can also employ a mouse lab or eye-tracking experiment to investigate consumer's actual information search and to investigate how restricting consumers to comparisons between certain well-defined reviews can affect their causal attributions.

Note

- 1 Two of the reviewer attribution items ("He doesn't know enough about tablet computers" and "He does not appear to have the expertise to evaluate the tablet computer properly") show a conceptual overlap with the expertise manipulation, threatening the discriminant validity between both variables. We therefore excluded these items from the analysis. Including the two scale items results in similar findings as the one presented in this paper. To be able to compare the results of Study 2 and Study 3 we re-ran Study 2 with the adjusted scale. All the findings reported in Study 2 were replicated in the re-analysis.

References

- Baumeister, F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323-370.
- Burnkrant, R. E., & Cousineau, A. (1975). Informational and normative social influence in buyer behavior. *Journal of Consumer Research*, 2(3), 206-215.
- Carlson, K. & Guha, A. (2010). The ratings paradox: Why we prefer reading negative reviews, but then subsequently rate these reviews as less useful. *Advances in Consumer Research*, 37.
- Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online word-of-mouth. *Journal of Marketing Research*, 50(4), 463-476.
- Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce*, 13(4), 9-38.
- Cheung, C. M. Y., Sia, C. L., & Kuan, K. K. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association for Information Systems*, 13(8). 618-635.
- Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3), 345-354.
- Dallal, G. V. (1999). *The little handbook of statistical practice*. Retrieved from <http://www.jerrydallal.com/LHSP/LHSP.HTM>
- Dawson, J. F. (2014). Moderation in management research: What, why, when and how. *Journal of Business and Psychology*, 29(1), 1-19.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3), 629-636.
- Doh, S. J., & Hwang, J. S. (2009). How consumers evaluate ewom (electronic word-of-mouth) messages. *CyberPsychology & Behavior*, 12(2), 193-197.
- Eastin, M. S. (2001). Credibility assessments of online health information: the effects of source expertise and knowledge of content. *Journal of Computer Mediated Communication*, 6(4).
- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38(6), 889-906.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York, NY: McGraw-Hill.
- Folkes, V. S. (1988). Recent attribution research in consumer behavior: A review and new directions. *Journal of Consumer Research*, 14(4), 13-23.
- Goldsmith, R. E., Lafferty, B. A., & Newell, S. J. (2000). The impact of corporate credibility and celebrity credibility on consumer reaction to advertisements and brands. *Journal of Advertising*, 29(3), 43-54.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Jiang, Z. J., & Benbasat, I. (2007). The effects of presentation formats and task complexity on online consumers' product understanding. *MIS Quarterly*, 31(3), 475-500.
- Kelley, H. H. (1967). *Attribution theory in social psychology*. In Nebraska Symposium on Motivation. University of Nebraska Press.
- Kelley, H. H. (1973). The process of causal attribution. *American Psychologist*, 28(2), 107-128.
- Laczniak, R. N., DeCarlo, T. E., & Ramaswami, S. N. (2001). Consumers' responses to negative word-of-mouth communication: An attribution theory perspective. *Journal of Consumer Psychology*, 11(1), 57-73.

- Lee, J., Park, D. H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, 7(3), 341-352.
- Littell, R. C., Stroup, W. W., & Freund, R. J. (2002). *SAS for linear models*. SAS Publishing, Cary, NC.
- Mizerski, R. W. (1982). An attribution explanation of the disproportionate influence of unfavorable information. *Journal of Consumer Research* 9(3), 301-310.
- Monga, A. B., & John, D. R. (2008). When does negative brand publicity hurt? The moderating influence of analytic versus holistic thinking. *Journal of Consumer Psychology*, 18(4), 320-332.
- Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, 34(1), 185-200.
- Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing*, 87(4), 598-612.
- Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, 1(1), 33-60.
- Purnawirawan, N., De Pelsmacker, P., & Dens, N. (2012). Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions. *Journal of Interactive Marketing*, 26(4), 244-255.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296-320.
- Russo, J. E., Meloy, M. G., & Medvec, V. H. (1998). Predecisional distortion of product information. *Journal of Marketing Research*, 5(4), 296-320.
- Schlosser, A. E. (2011). Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments. *Journal of Consumer Psychology*, 21(3), 226-239.
- Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing*, 21(4), 76-94.
- Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52(4), 689-699.
- Willemsen, L. M., Neijens, P. C., Bronner, F., & de Ridder, J. A. (2011). "Highly recommended!" The content characteristics and perceived usefulness of online consumer reviews. *Journal of Computer-Mediated Communication*, 17(1), 19-38.
- Willemsen, L. M., Neijens, P. C., & Bronner, F. (2012). The ironic effect of source identification on the perceived credibility of online product reviewers. *Journal of Computer-Mediated Communication*, 18(1), 16-31.
- Wu, P. F. (2013). In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology & Marketing*, 30(11), 971-984.
- Yang, J., & Mai, E. (2010). Experiential goods with network externalities effects: An empirical study if online rating system. *Journal of Business Research*, 63(9-10), 1050-1057.

About the Authors

Simon Quaschnig (Ph.D., Ghent University) is a postdoctoral researcher in Marketing at Ghent University (Belgium). His academic interests focus on the use of information in online decision-making.

Address: Department of Marketing, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium.
E-mail: simon.quaschnig@ugent.be

Mario Pandelaere (Ph.D., KU Leuven) is a Professor in Marketing at Ghent University (Belgium). His research interests mainly focus on social influence and persuasion, quantitative information processing and materialistic consumption.

Address: Department of Marketing, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium.
E-mail: mario.pandelaere@ugent.be

Iris Vermeir (Ph.D., Ghent University) is an Associate Professor in Marketing at Ghent University (Belgium). Her research interests mainly focus on attitudes and behavior, personality influences on consumption and consumer ethics.

Address: Department of Marketing, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium.
E-mail: iris.vermeir@ugent.be

Appendix 1

Example of the stimuli with the third review as the focal review (positive focal review in a positive context)



Sarah Vermeulen, 25

I visited the restaurant together with friends and it was disappointing. It was very uninviting, the interior was ugly, the food wasn't good and the service was bad.



Tom Peeters, 24

I had recently eaten here and was surprised how good it was. The restaurant was very busy, but the staff was very friendly. I had the lasagna, which was just delightful.



Pieter Janssens, 23

I had a fantastic experience. Good service, great food, large portions and an intimate and cozy atmosphere. What more can you ask for? The price was higher than I would normally pay, but it was totally worth it.



Elke Van Doren, 27

I really enjoyed the food. This is a good Italian option in this neighborhood. The gnocchi was not the best, but everything else was very tasty. I would definitely recommend this place!
