Taking a deeper look at online reviews:

The asymmetric effect of valence intensity on shopping behaviour

Accepted for Publication in the Journal of Marketing Management

Dr. Arne Floh (first author, corresponding author)
Surrey Business School
Guildford, GU2 7XH, Surrey, UK
Tel: +44 (1483) 689185
E-mail: a.floh@surrey.ac.uk

Dr. Monika Koller
WU Vienna (Vienna University of Economics and Business)
Department of Marketing
Augasse 2-6, 1090 Vienna, Austria
Tel.: +43 (1) 31336-5330
E-mail: monika.koller@wu.ac.at

Dr. Alexander Zauner
WU Vienna (Vienna University of Economics and Business),
Department of Marketing
Augasse 2-6, 1090 Vienna, Austria
Tel: +43 (1) 31336-4098
E-mail: alexander.zauner@wu.ac.at
Taking a deeper look at online reviews:

The asymmetric effect of valence intensity on shopping behaviour

Abstract

This study tests the asymmetric effect of user-generated, open-ended online reviews on online shopping behaviour (intention-to-buy, intention-to-recommend, and willingness-to-pay). Three online experiments involving manipulating the valence intensity of online reviews for hotels, books, and running shoes (overall customer sample of n=818) provide empirical support for the proposed relationship. The valence intensity of online reviews moderates the effect of online reviews on purchase intentions. In other words, a significant change in online shopping behaviour was found for positive medium and strong reviews, but not for negative ones. Based on these findings, managers should encourage customers to share their positive consumption-related experiences by offering strong arguments that will convince other customers.

Summary statement of contribution (75 words max):

The present paper contributes to the research on the effect of online reviews on online shopping behaviour. Based on the findings of three online experiments, it can be concluded that valence intensity moderates the relationship between online reviews and online shopping behaviour. The occurrence of slightly negative information chunks in online reviews can cause significant negative changes in online shopping behaviour.

Keywords: Electronic Word-of-Mouth, Online Reviews, Valence Intensity, Online Experiment
Taking a deeper look at online reviews:

The asymmetric effect of valence intensity on shopping behaviour

The rapid diffusion of the internet has provided consumers with an easy and frequently used way to compare different market offerings and search for consumption-related advice given by other consumers in the form of electronic word-of-mouth (eWoM) (Henning-Thurau & Walsh, 2003; Park & Lee, 2009). For example, Gretzel and Yoo (2008) report in their study that more than 74% of travellers use the comments of other customers as an information and decision source when planning trips for pleasure. According to the same study, 44% of all internet users in the US have engaged in eWoM activities at least once.

EWoM encompasses a plethora of different forms and technologies (e.g. blogs, tweets and shopping bots) of which online ratings and online reviews are the most accessible and most frequently used (Trusov, Bucklin, & Pauwels, 2009). Online ratings are a quantitative summary of experiences, attitudes and opinions, usually expressed as stars or points, whereas reviews are open-ended, user-generated text messages about a product or service (Park & Kim, 2008; Mudambi & Schuff, 2010). The two forms can occur separately or in combination. The most prominent example of online ratings and online reviews is that by online retailer Amazon.com.

Previous research has primarily focused on the average online rating and the number of online reviews (e.g. Zhu & Zhang, 2009). The variability of online review texts has been ignored or the complexity has been reduced to a simple dichotomous case (positive versus negative reviews) (see Table 1 for an overview of related studies). This is unfortunate for at least two reasons: First, it is more common to find reviews including both positive and negative facets simultaneously. Second, given the huge number of online reviews available (for example, in June 2012, TripAdvisor contained more than 75 million online reviews), a simple dichotomisation simply does not reflect reality.
The present experimental study is motivated by Cheung, Luo, Sia, and Chen (2009), who suggested examining positively versus negatively framed reviews (relating to the perceived valence of an online review) and one-sided versus two-sided reviews (relating to the perceived valence intensity of an online review text) in more detail. Hence, this study aims at empirically testing whether the perceived valence intensity of online reviews has an impact on online shopping behaviour (intention-to-buy, intention-to-recommend, and willingness-to-pay). Overall, 818 customers participated in three online experimental studies. The valence intensities of the presented online reviews were manipulated in order to study the effect on online shopping behaviour regarding three different products (books, hotels and running shoes).

This study contributes to the existing eWoM literature in the following ways: First, to the best of our knowledge, we are not aware of any study analysing the impact of the valence intensity of online reviews on online shopping behaviour. Whereas variations in online ratings have been researched in the past, we aim to fill the gap for online reviews. Second, we extend the existing research on the valence of online reviews. Previous research has applied single-student samples from the US, Asia and Australia using books and hotels. We have used a more comprehensive approach (three studies and three different products) with customers from Europe. Third and finally, our study can be linked to the research into eWoM effects on online shopping behaviour. We extend previous research by covering three different aspects of online shopping behaviour: intention-to-purchase, intention-to-recommend, and willingness-to-pay.

The rest of the paper is organised as follows: We start with a literature review focusing on online ratings and online reviews, followed by a discussion of the theoretical background and the research hypotheses. The methods and results of the three experimental studies are presented next. The paper concludes with a discussion of the results, suggestions for marketing research and practice, limitations and ideas for future research.
Table 1 gives an overview of selected studies on eWoM. Studies were included based on the following two criteria: i) Only studies about online reviews and online ratings were selected. Prior research on traditional WoM or other forms of eWoM were not included. ii) Only studies dealing with online shopping behaviour as an outcome of eWoM are included. Studies with output variables such as information credibility or the helpfulness of reviews (e.g. Mudambi & Schuff, 2010) are not listed in the table. The following conclusions can be drawn from Table 1:

i) Overall, the impact of eWoM on online shopping behaviour has been empirically validated. Average online ratings and the number of online reviews positively influence purchase intentions, hotel bookings, and sales.

ii) The average rating and the number of reviews are the research focus for most of the studies. The valence (sometimes referred to as quality) of online reviews has rarely been researched. This is surprising given the fact that online review texts are considered to be more impactful compared to online ratings. These differences in research frequencies may be explained by the easier accessibility of online ratings and the increased effort necessary to analyse open-ended, user-generated review texts.

iii) Mixed results were found analysing the variance of ratings. Whereas, a study on restaurants supports the positive effect on purchase intentions, two studies with contradicting results were found in the hotel industry. To the best of our knowledge, no study was found dealing with the valence intensity of online reviews.

iv) In terms of study characteristics, most of the studies were conducted in the US and China using secondary data collected from online retailers or online recommendation websites. None of the studies were conducted in Europe. Hotels, movies and books are the most frequently analysed products. Student sampling was the most common method used in experiments.
Table 1: Selected studies examining the effect of online ratings and online reviews on online shopping behaviour

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Product</th>
<th>Data/ Method</th>
<th>Country</th>
<th>Sample Size</th>
<th>Independent Variable(s)</th>
<th>Dependent Variable</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen et al. (2004)</td>
<td>Book</td>
<td>Secondary</td>
<td>US</td>
<td>610</td>
<td>Average rating, number of ratings</td>
<td>Number of reviews</td>
<td>Sales</td>
</tr>
<tr>
<td>Chevalier/Mayzlin (2006)</td>
<td>Book</td>
<td>Secondary</td>
<td>US</td>
<td>1,636</td>
<td>Average rating</td>
<td>Number and length of reviews</td>
<td>Sales</td>
</tr>
<tr>
<td>Dellarocas et al. (2007)</td>
<td>Book</td>
<td>Secondary</td>
<td>US</td>
<td>55,156</td>
<td>Average rating</td>
<td>Number of reviews</td>
<td>Sales</td>
</tr>
<tr>
<td>Duan et al. (2008)</td>
<td>Movie</td>
<td>Secondary</td>
<td>US</td>
<td>64</td>
<td>Average rating</td>
<td>Number of reviews</td>
<td>Sales</td>
</tr>
<tr>
<td>Park/Lee (2009)</td>
<td>Book/ School Programme</td>
<td>Experiment</td>
<td>US</td>
<td>440*</td>
<td>Valence of reviews</td>
<td>Influence on purchase decision</td>
<td>Significant effect of valence, negative review texts have a higher impact on purchase decisions</td>
</tr>
<tr>
<td>Park/Lee/Han (2007)</td>
<td>Multimedia</td>
<td>Experiment</td>
<td>NA</td>
<td>352*</td>
<td>Number and helpfulness of reviews</td>
<td>Purchase intention</td>
<td>Significant effect of number and helpfulness of reviews</td>
</tr>
<tr>
<td>Park/Kim (2008)</td>
<td>Multimedia</td>
<td>Experiment</td>
<td>NA</td>
<td>222*</td>
<td>Number of reviews</td>
<td>Purchase intention</td>
<td>Significant effect of number of reviews on purchase intention</td>
</tr>
<tr>
<td>Sparks/Browning (2011)</td>
<td>Hotel</td>
<td>Experiment</td>
<td>Australia</td>
<td>554</td>
<td>Valence of reviews, Order of reviews</td>
<td>Purchase intention</td>
<td>Significant effect of valence and order of reviews (negative reviews have a higher impact)</td>
</tr>
<tr>
<td>Ye et al. (2009)</td>
<td>Hotel</td>
<td>Secondary</td>
<td>China</td>
<td>3,625</td>
<td>Average and variance of ratings</td>
<td>Sales</td>
<td>Significant effect of mean and variance of ratings</td>
</tr>
<tr>
<td>Ye et al. (2011)</td>
<td>Hotel</td>
<td>China</td>
<td></td>
<td></td>
<td>Average and variance of ratings</td>
<td>Number of bookings</td>
<td>Significant effect of average rating. Non-significant effect of variance of ratings on bookings</td>
</tr>
<tr>
<td>Zhang et al. (2010)</td>
<td>Restaurants</td>
<td>Secondary</td>
<td>China</td>
<td>1,242</td>
<td>Average rating</td>
<td>Number of reviews</td>
<td>Website popularity</td>
</tr>
<tr>
<td>Zhu/Zhang (2010)</td>
<td>Games and consoles</td>
<td>Secondary</td>
<td>US</td>
<td>3,330</td>
<td>Average and variance of ratings</td>
<td>Number of reviews</td>
<td>Sales</td>
</tr>
</tbody>
</table>

Note: NA = Not available; * = Student sampling; † = Valence of reviews was measured with two categories only (binary coding: negative/positive)
Theoretical foundations of hypotheses

eWoM influences online shoppers if they base their purchase intentions and/or behaviour on online reviews (Senecal & Nantel, 2004). Aside from possible moderating factors, such as website reputation (Park & Lee, 2009), the effect depends greatly on the message itself. In personal WOM, an important dimension of a WOM message is the strength of delivery (delivering the message with strong words or powerfully) (Mazzarol, Sweeney, & Soutar, 2007). In written WOM messages, such as online reviews, strength of delivery is reflected in the perceived valence intensity of the review text. The argument strength is the extent to which the receiver of the message views the argument as valid, convincing and reinforcing its position (Cheung et al., 2009). The arguments given in an online review represent the informational part of the message, which also carries an affective part, eliciting a motivational component. This motivational component leads to the eWOM effect (either following the suggestion or disregarding it).

Following the circumplex model of Russel (1980), an affect consists of a valence and an arousal dimension. Hence, an online review can also be understood as carrying a valence part (positive versus negative) and an arousal in terms of an intensity part (e.g., high versus only moderate). The two parts are interlinked. Since, in the case of online reviews, the valence and the intensity appear conjointly (e.g., highly positive/negative or only moderately so or neutral), the term “valence intensity” is used in the present study to describe the degree of positivity or negativity, or alternatively the neutrality of the review content.

Previous research on eWoM has mainly been confined to how the valence of the reviews (Lee & Youn, 2009; Park & Lee, 2009; Tsang & Prendergast, 2009), information credibility (Cheung et al., 2009) and the appropriate number of reviews (Park, Lee, & Han, 2007) explain purchase intentions. While the literature suggests that scholars have already agreed upon the
general existence of an eWoM effect on purchase intentions (see, e.g., Park & Lee, 2009; Park & Kim, 2008), conflicting results have been found as to the effects of positive versus negative product reviews (Arndt, 1967; East, Hammond, & Lomax, 2008; Park & Lee, 2009). To solve this issue, some scholars have suggested testing for moderating variables, such as gender (Park, Yoon, & Lee, 2009), product knowledge, involvement (Lee, Park, & Han, 2008), interpersonal differences (Chen, Shang, & Kao, 2008) or pre-WOM product or brand preferences (East et al., 2008), to explain the diversity of previous research findings.

However, another possible explanation for these inconsistencies could lie in the mere dichotomisation of reviews into either positive or negative ones, disregarding the valence intensity. Such a simple dichotomisation of consumer reviews is not grounded in reality. In fact, product reviews including purely positive or negative information are hard to find. More often, product reviews contain a mix of positive and negative product information chunks. This mix of information leads to a certain level of valence intensity perceived by the eWoM receiver, depending on whether the negative or the positive information chunks contribute more to their overall perception of the online review.

The overall effect of valence intensity

The hypothesis that the valence intensity of an online review has a significant effect on behavioural intentions is theoretically grounded on the Hull-Spence Behavior Theory (HSBT). Based on Hull (Hull, 1943, 1958; Hull et al., 1940) and expansions of his theory by Spence (1956, 1960), we assume that strong, purely positive or purely negative arguments (high valence intensity) have a stronger impact on behavioural intentions than messages with mixed positive and negative content (resulting in a valence intensity level that is perceived as medium or moderate). Originally introduced as reinforcement theory, Hull attempted to construct a theory on
the basis of the hypothetico-deductive approach in an algebraic form to evaluate and understand all behaviour, ranging from the simple behaviours of laboratory animals to purposeful human behaviours.

There have been several attempts to promote Hull’s theory as a paradigm for the study of consumer behaviour. For example, Hawkins (1970) and Kivetz, Urminsky, and Zheng (2006) have successfully applied and empirically tested the HSBT in the context of advertising research. Whereas the former focused on the effects of subliminal stimulation, the latter analysed certain input and output variables, such as delay of reward and response speed, with conflicting results. Recently, Rossiter and Foxall (2008) have even argued that the HSBT constitutes an ‘all-encompassing theory of the type that was popular initially in consumer behaviour’.

This psychological theory by Hull (1943) has been used to causally relate, not only independent variables (e.g., stimulus intensity dynamism) and dependent variables (e.g., amplitude or speed), but also mediating and moderating variables (e.g., reaction potential) that were causally related. The independent variable stimulus intensity dynamism, in particular, is of critical relevance for this study. This variable indicates that a stronger stimulus will produce a stronger response, holding all other variables constant (Ward & Robertson, 1973). In other words, a strong, purely positive or purely negative eWoM stimulus will increase the propensity to respond and subsequently lead to a stronger response in terms of an enhanced purchase intention in the case of strong positive and a significantly reduced purchase intention in the case of strong negative reviews. In the case of only moderately positive or negative reviews, the effect on purchase intentions is expected to be less pronounced.

This assumed effect of valence intensity derived from HSBT is also backed up by findings on the extremity effect. The extremity effect describes the notion that information receivers generally give extreme messages more weight (Fiske, 1980). Derived from theoretical
reasoning based on HSBT and the extremity effect, as well as from findings by Lee et al. (2009) in the context of eWoM extremity and its effect on brand attitude, an online review of high valence intensity (in terms of strong arguments) is assumed to be more effective at altering actual behavioural intentions than a text of medium intensity.

Based on this theoretical background, we state the first of our two main research hypotheses:

H1: The valence intensity of an online review is positively associated with the intention to purchase a product.

*The asymmetric effect of valence intensity*

The general issue of whether positive or negative eWoM has a greater impact has received considerable attention (e.g., East et al., 2008; Park & Lee, 2009). Moreover, in cases where single reviews for a particular product are inconsistent (some positive, others negative), Tsang and Prendergast (2009) report that the so-called negativity effect is the underlying force that determines which reviews consumers rely on more. The negativity effect implies that positive and negative information do not necessarily have a symmetrical influence on consumers’ decision making. Rather, the harmful impact of negative information is much greater than the beneficial impact of positive information (Tsang & Prendergast, 2009). The psychological literature on the negativity effect describes this mechanism as being due to the greater power of bad events compared to that of good ones (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Peeters & Czapinski, 1990; East et al., 2008). Peoples’ tendency to give negative information more attention is also grounded in evolution, as people tended to be more likely to survive if they tried to avoid negative contacts and consequences (Fiske, 1980).
This psychological mechanism underlying the negativity effect is also substantiated in prospect theory. In general, prospect theory suggests that people are risk averse in conditions where they expect to gain something, and risk seeking in conditions where they stand to lose something (Kahneman & Tversky, 1984). In the context of purchase decision making, individuals generally face a reward-seeking condition; they want to gain something in terms of finding a good product that will deliver them multidimensional value and satisfy their needs. Following this reasoning, if people read online reviews prior to making a purchase decision, they might be more alert to any negative information given in the online reviews as they will be trying to avoid risk as much as possible.

Based on the effect of negativity embedded in the frame of prospect theory, the eWOM effect of strong positive reviews should be reduced as soon as negative information chunks appear in the picture. Just a few chunks of negative information might increase the perceived risk in what is generally a gain-seeking condition. Hence, the individual reading an only medium-positive review might react to the small number of negative aspects and be significantly less prone to purchase the product. However, the valence intensity of the online review is expected to have an impact only in the case of generally positive reviews, and not in the case of negative ones. This assumption is again backed up by the negativity effect. The slightest chunk of negative information looms larger than any positive chunks and significantly lowers the purchase intention immediately. Hence, it should make virtually no difference whether a review is only moderately negative or highly negative. In both cases, the purchase intention is expected to be significantly reduced.

Based on this theoretical reasoning, the second research hypothesis is as follows: $H_2$: The relationship between the valence intensity of a review and purchase intentions is asymmetric.
Study 1: Student sample

Method

Participants and design. The online experiment method was chosen since the test situation equates to a ‘real’ product evaluation and purchase decision-making process in e-commerce, based on online reviews. In this case, choosing the internet as the setting for the experiment enhances its external validity (for a methodological review, see Reips, 2002a, 2002b, 2002c).

An electronic invitation to take part in the research project was emailed to a student mailing list of a large European business school. As many students buy books or book hotels (the products studied in this experiment) using the internet, a student sample qualifies as being applicable. Five vouchers for a travel agency were promised to five randomly drawn students as an incentive. The experiment employed a 2 (valence: positive, negative) x 2 (valence intensity: medium, strong) x 2 (product: book, hotel) within-subject design. Overall, 339 respondents or 15.1 (211 female, 128 male; mean age 25 years) completed the online experiment.

Stimuli. Books and hotels were chosen as the objects of investigation in order to make a comparison between search (books) and experience (hotels) goods. These goods were chosen for this study of buying decisions via the internet because books and hotels are two of the major product categories sold over the internet and the majority of consumers are familiar with both. Two pictures (one book cover and one picture of a hotel) served as product stimuli. Four different online reviews (positive or negative and medium or high valence intensity) were used as manipulative stimuli for each product type. Each review comprised subjective evaluations, product information and stories about experiences with the book or hotel. The product stimuli, as well as the valences and valence intensities of the online reviews, were pre-tested in an online experiment (n=56). Additionally, technical issues, such as the randomised assignment of respondents and data exporting, were tested before the main study was launched.
Procedure. The cover story explained that the study was about shopping behaviour on the internet. The participants were informed that the study was anonymous. However, an email address had to be provided if they wanted to take part in the lottery draw for the travel agency vouchers. After answering some demographic questions, the participants went through different phases of the experiment in the following order: conditioning phase (book), manipulation and control check (book), purchase intention measurement (book), conditioning phase (hotel), manipulation and control check (hotel), purchase intention measurement (hotel). The participants were then thanked and reminded to provide an email address for the lottery draw.

Measurement. Purchase intention, measured by the items ‘How likely is it that you will buy the book?’ and ‘How likely are you to book a room in this hotel?’, respectively, was chosen as the dependent variable. A five-point rating scale, with end-points verbalised as 1 = ‘very unlikely’ and 5 = ‘very likely’, was employed.

Results

Manipulation check. As a first step, the authors tested the desired effects of the manipulative stimuli. For this purpose, the respondents were asked about how they perceived the quality of the online reviews for each manipulation group separately. A semantic profile ranging from negative to positive was used to measure these items (five-point Likert scale). If the manipulated online reviews performed as intended, the authors would expect to see a significant and positive correlation between the manipulation variable, namely the perceived valence intensity of the online reviews, and the dependent variable, namely purchase intention. A series of Pearson correlation tests revealed significant and substantial coefficients ranging from 0.174 to 0.655 (all p <0.05). Based on these results, we conclude that the quality of the manipulated online reviews was perceived by the respondents as intended by the researchers.
Halo and memory effect. The respondents had to accomplish four different tasks consecutively. They were randomly assigned to each of the four experimental or control conditions in turn, so that the order of completion differed among the participants. The authors conducted a robustness check to evaluate whether the sequence in which the participants completed the tasks influenced their responses. For this purpose, the corrected purchase intentions of conditions #1 and #3 (books) were correlated with the corrected purchase intentions of conditions #2 and #4. None of the inter-associations were significant (all p >0.17), suggesting there was no bias in the results due to a mere measurement effect such as the halo and/or the memory effect (Lachman & Bass, 1985; Thorndike, 1920).

Valence intensity. H₁ proposes that the valence intensity of online reviews has a positive significant influence on a customer’s purchase decision. The data for the positive and negative reviews were aggregated, yielding two new variables for medium and high valence intensity. Afterwards, a paired sample t-test was conducted. The results show a significant effect of valence intensity for books ($M_{\text{medium}} = 0.80$, $M_{\text{high}} = 1.08$, $t (1, 138) = -3.381; p <0.01$) and hotels ($M_{\text{medium}} = 0.79$, $M_{\text{high}} = 1.13$, $t (1, 153) = -7.175, p <0.01$). Hence, H₁ is supported by the data.

Assymetric effect of valence intensity. H₂ deals with the asymmetric effects of the valence intensity of reviews. It proposes that valence intensity has a varying influence on purchase intentions, depending on the direction of the review. Therefore, two ANOVAs were conducted for the dependent samples (book and hotel). The results are shown in Figure 1.
Figure 1: Results of Study 1

Hypothesis 1

![Graph showing the results of Study 1 for Hotel Book]

The findings support the proposed asymmetric influence of valence intensity on purchase intentions, affirming H2. Whereas strong (high valence intensity) positive reviews have a significantly stronger effect on purchase intentions than medium (valence intensity) positive reviews, the same effect is not significant for negative reviews. In other words, the strength of negative eWOM does not matter for consumer decision making. Consistently, this result is true for both the book (M_{positive - medium} = 0.84, M_{positive - high} = 1.32, F(1, 40) = 9.962, p < 0.01; M_{negative - medium} = 0.80, M_{negative - high} = 0.80, F(1, 37) = 0.001, p > 0.1) and the hotel reviews (M_{positive - medium} = 0.84, M_{positive - high} = 1.32, F(1, 40) = 9.962, p < 0.01; M_{negative - medium} = 0.80, M_{negative - high} = 0.80, F(1, 37) = 0.001, p > 0.1).
Discussion

This study served as a pilot for testing the influence of the valence intensity of online reviews on purchase intentions. Using a large student sample, the authors assessed the internal validity of the proposed relationships. As expected, the basic hypotheses are supported by the data. Messages with stronger valence intensity lead to higher levels of purchase intentions. However, this effect was only observed in the case of positive messages, supporting the concept of the negativity effect (Tsang & Prendergast, 2009).

In the next step, the external validity of the results will be checked. For this purpose, a second study is conducted with the following adaptations made: i) The setting is broadened to ‘realistic’ verbal (online reviews) and graphical (pictures) stimuli for the products. ii) A random sample drawn from the customers of an online retailer is used. iii) Demographic variables and perceived message credibility are added as control variables in order to check for any confounding effects.

Perceived credibility is defined as the extent to which the reader perceives an online review to be believable, true or factual (Cheung et al., 2009). Applying the persuasion effect to eWOM, it is proposed that mixed reviews (reviews containing positive and negative information chunks) are perceived as more credible by customers, whereas pure reviews (either positive or negative) are regarded as less trustworthy. Hence, the asymmetric effect of valence intensity could be confounded by the perceived credibility of the review.
Study 2: Customer sample

Method

Participants and design. This experiment employed a 2 (valence: positive, negative) x 2 (valence intensity: medium, strong) x 2 (product: book, hotel) mixed design with the first two factors manipulated between subjects and the latter factor manipulated within subject. An email invitation was sent to a customer listing of an online retailer. Again, each respondent was randomly assigned to an experimental group. Table 2 shows the characteristics of the sample. The final sample includes 317 customers (completion rate 81.7 %).

Table 2: Sample Characteristics of Study 2

<table>
<thead>
<tr>
<th>Gender</th>
<th>Frequency</th>
<th>Education</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>165 (52.1)</td>
<td>Minimum Compulsory School</td>
<td>49 (15.5)</td>
</tr>
<tr>
<td>Female</td>
<td>152 (47.9)</td>
<td>High School</td>
<td>134 (42.3)</td>
</tr>
<tr>
<td>Age (Mean)</td>
<td>35</td>
<td>College</td>
<td>134 (42.3)</td>
</tr>
</tbody>
</table>

Stimuli. Originally, eight online reviews and four pictures for each product category were selected from Amazon and the website holidaycheck to serve as stimuli. In an online pre-test (n=30), respondents rated the reviews and pictures on a seven-point rating scale (positive to negative). Based on these scores, we chose online reviews with a variety of scores (strong positive, medium positive, medium negative, and strong negative) for valence intensity and picture stimuli with neutral scores (3.5 for book, 3.6 for hotel) to include in the study.

Procedure and conditioning phase. The conditioning and measurement procedure was analogous to that used in Study 1. Participants were randomly assigned to one of the experimental groups and shown a common picture stimulus and manipulated reviews. Each review comprised subjective evaluations, product information and stories about an experience with the book or hotel. The participants went through the experiment in the following order: conditioning phase (book), manipulation and control check (book), purchase intention.
Measurement. The same items were used as in the first study to measure the dependent variable (purchase intention). Additional items were used to measure message credibility (Cheung et al., 2009). A full list of items is shown in Appendix A.

Results

Manipulation check. The respondents were asked about how they perceived the valence intensity of the online review (1 = positive, 5 = negative). An ANOVA analysis showed significant differences between the four online reviews (strong positive, medium positive, medium negative, and strong negative) for both products (hotel: F(3, 316) = 220.48, p = 0.000; book: F(3, 316) = 220.48, p = 0.000). Figure 2 shows a bar chart of the manipulation check.

Valence intensity. Aggregated means of the medium and high valence intensity reviews are shown in Figure 2. Consistent with H₁, two independent t-tests revealed significant differences in purchase intentions due to the valence intensities of the online reviews for the book (M_{medium} = 1.87, M_{high} = 2.21, t (1, 181) = -2.060, p <0.05) and the hotel (M_{medium} = 1.69; M_{high} = 2.06, t (1, 189) = -2.390, p <0.05).

Asymmetric effect of valence intensity. The findings of Study 2 also support the proposed asymmetric influence of valence intensity on purchase intentions. Whereas strong (high valence intensity) positive reviews have a significantly stronger effect on purchase intentions than medium (valence intensity) positive reviews, the same effect is not significant for negative online reviews. Again, this result holds for both the book (M_{positive - medium} = 2.00, M_{positive - high} = 3.27; t (1, 86) = 5.049, p <0.001; M_{negative - medium} = 1.69, M_{negative - high} = 1.53, t (1, 95) = -1.033; p >0.3)
and the hotel reviews ($M_{\text{positive - medium}} = 2.08, M_{\text{positive - high}} = 3.47; t(1, 100) = 6.690; p < 0.001$

$M_{\text{negative - medium}} = 1.29; M_{\text{negative - high}} = 1.19, t(1, 125) = -1.0707; p > 0.2$).

**Figure 2: Results of Study 2**

**Manipulation Check**

**Hotel**

- **Positive High Valence Intensity**: 4.50
- **Positive Medium Valence Intensity**: 2.98
- **Negative High Valence Intensity**: 2.09
- **Negative Medium Valence Intensity**: 1.23

**Book**

- **Positive High Valence Intensity**: 4.70
- **Positive Medium Valence Intensity**: 3.34
- **Negative High Valence Intensity**: 2.53
- **Negative Medium Valence Intensity**: 1.61

**Hypothesis 1**

**Purchase Intention**

- **Book**: 2.21, 1.69, 2.08
- **Hotel**: 1.87

**Hypothesis 2**

**Hotel**

- **Positive**: 3.47, 2.09, 1.28, 1.18
- **Negative**: 1.19

**Book**

- **Positive**: 3.27, 2.00, 1.08, 1.53
- **Negative**: 1.53
Message credibility. Message credibility was tested to see whether it had a confounding effect on the relationship between the valence intensity and purchase intentions. The credibility of the eWOM messages were found to have a significant effect on purchase intentions (hotel: $F(1, 181) = 5.067, p<0.05$; book : $F(1, 125) = 40.365, p <0.001$). However, the main finding of our study (the asymmetric effect of valence intensity) is not confounded by the perceived credibility. Valence intensity remains significant in the conditions with positive reviews only (all $p <0.001$) and is not significant in the negative conditions (all $p >0.3$).

Demographic control variables. The second study was also used to test for the effect of demographics (gender, education and age). The results confirm the validity of the initial results. Whereas gender and education have no significant effect on purchase intentions ($p >0.4$), age does ($F(1,316) = 9.618, p = 0.002$). However, the proposed asymmetric effect is not altered by age. Valence intensity remains significant in the conditions with positive reviews only (all $p<0.001$) and is not significant in the negative conditions (all $p >0.3$).

Discussion

The aim of the second study was to increase the validity of the results obtained in Study 1. For this purpose, realistic picture stimuli and online reviews were used. Additionally, a random customer sample was drawn from an online retailer. Demographics (age, gender, and education) and perceived message credibility were used as control variables. Nevertheless, the results remain the same. Positive eWOM messages with a stronger valence intensity lead to higher levels of purchase intentions. However, the opposite is not true for negative messages.

Study 3 was conducted using three categories of valence intensity (positive, neutral and negative). We reduced the number of categories to further enhance the validity of our results found in Studies 1 and 2. Confirming the results with a smaller number of categories would be a
strong robustness test of the asymmetric effect of valence intensity. Based on the theoretical role of risk aversion in gain-seeking conditions, as described in prospect theory, which is the one of the underlying psychological underpinnings guiding our study, the perceived risk of online shopping was included as a new control variable. Product involvement was also included, since involvement in a specific product category might also impact on how an individual deals with negative chunks of information. Moreover, we extended the range of dependent variables to tap even more into relevant real-life purchase-related phenomena. We added willingness-to-pay and intention-to-recommend to the original purchase intention variable. Additionally, a new product (running shoes) was chosen to test for the proposed effects, adding further to the generalisability of the effect of valance intensity across various product categories.

Study 3: Student sample

Method

Participants and design. This experiment employed a one-factor (positive, neutral or negative), between-subject design (product: running shoes). Again, an online experiment was used. Students were chosen as the participants since no significant differences in results were found between Studies 1 and 2. Again, a mailing list of a large European Business School was used for recruiting participants. 162 (93 female, 71 male; mean age: 25 years) students participated in Study 3 (completion rate 7.2 %).

Stimuli, procedure and measurement. Again, real-world online reviews and pictures were chosen and pre-tested. In contrast to Studies 1 and 2, only three categories of valence intensity were used (positive, neutral and negative). The conditioning and measurement procedure was analogous to that used in the first two studies. Additional control variables (perceived risk and product involvement) and dependent variables (willingness-to-pay and intention-to-recommend)
were used, as explained in the last section. Established scales were used to measure the new control and dependent variables: involvement (Bloch, 1981), perceived online risk (Forsythe, Liu, Shannon, & Gardner, 2006), willingness-to-recommend (Zeithaml, Berry, & Parasuraman, 1996), and willingness-to-pay (Homburg, Koschate, & Hoyer, 2005). A full list of items is shown in Appendix A.

Results

Manipulation check. Again, respondents were asked about the perceived valence intensity of the online reviews (1 = positive, 5 = negative). An ANOVA analysis showed a significant overall effect (F(2, 161) = 447.42, p = 0.000). A subsequent post hoc test confirmed the significant differences across the three online reviews (positive, neutral and negative).

Assymetric effect of valence intensity. The findings of Study 3 support the proposed asymmetric influence of valence intensity on purchase intentions (overall effect (F(2, 161) = 22.82, p = 0.000)). Whereas positive reviews have a significantly stronger effect on purchase intentions than neutral reviews (M\text{positive} = 2.93, M\text{neutral} = 1.93; t (1, 108) = 5.081; p <0.001, no difference in effect sizes was found between the negative and neutral reviews (M\text{negative} = 2.93, M\text{neutral} = 1.85; t (1, 96) = -0.514; p >0.1. Similar results were found for the two additional dependent variables – willingness-to-pay (overall effect (F(2, 161) = 24.487, p = 0.000; M\text{positive} = 71.38, M\text{neutral} = 55.11; t (1, 108) = 2.495; p <0.05; M\text{negative} = 43.67, M\text{neutral} = 55.11; t (1, 96) = -0.824; p >0.1; M\text{positive} = 71.38, M\text{negative} = 1.93; t (1, 114) = 4.559; p <0.001) and intention-to-recommend the product (overall effect (F(2, 161) = 10.658, p = 0.000; M\text{positive} = 2.70, M\text{neutral} = 1.74; t (1, 96) =5.126; p >0.1; M\text{negative} = 1.46, M\text{neutral} = 1.74; t (1, 86) = -0.682; M\text{positive} = 2.70, M\text{negative} = 1.46; t (1, 114) = 7.557; p <0.001). Figure 3 shows the bar charts for all dependent variables.
Control variables. The third study was also used to test for the effect of additional control variables. The results confirm the validity of the initial results. Whereas product involvement has no significant effect on purchase intentions ($F(1, 161) = 0.016; p = 0.898$), perceived risk of buying online does ($F(1, 161) = 9.34; p = 0.002$). However, the proposed asymmetric effect is not altered by perceived risk. Valence intensity remains significant in the conditions with positive reviews only (all $p <0.001$) and not significant in the negative conditions (all $p >0.2$).
Discussion

The aim of the third study was to further validate the findings. For this purpose, a new product and just three categories of valence intensity (instead of four) were chosen. By replicating the results in a new experimental setting, this paper makes a stronger contribution. Product involvement and perceived risk do not alter the asymmetric effect. Interestingly, the effect can also be found when variables such as willingness-to-pay and intention-to-recommend are used as dependent variables.

Conclusion and outlook

Research implications

WoM communication has a long tradition in marketing research. The research stream has become even more popular and important since the emergence and rapid diffusion of the internet. Many empirical studies have provided evidence of the effect of eWOM in general. However, as we stated in the theoretical part of this paper, findings about the valence and valence intensity of online reviews are limited. Furthermore, previous findings about the direction of reviews are somewhat inconsistent and can be questioned over their simplification of the data into either purely positive or purely negative reviews. In fact, more frequently, a combination of both positive and negative characteristics are found within product reviews. We target the valence of product reviews in general, and the valence intensity in particular, in this study.

The findings of the experiments support our hypotheses. Purchase intention is largely influenced by the valence intensity of online reviews. Interestingly, the analyses reveal an asymmetric effect. A significant increase in purchase intentions between medium and strong reviews was found for the positive reviews only. These results also held when controlling for
demographics, the perceived information credibility of the message, product involvement and perceived risk.

Additionally, the results are robust across books, hotels and running shoes. This finding contradicts the results of prior research (e.g. Park & Lee, 2009). A potential explanation could be the huge amount of information that is now available on the internet. In June 2012, more than 75 million online reviews were available on TripAdvisor. Similarly, Duan, Gu and Whinston (2008b) report an average of 1,350 online reviews per movie. Based on these impressive numbers and the rapid dispersion of online reviews, we assume that there will gradually be a declining difference between the search and experience attributes of products.

From a psychological point of view, people tend to strive to maintain consistency in everyday life in general and in consumption situations in particular (Cialdini, Trost, & Newsom, 1995; Festinger, 1957). In online consumption decisions, reading consumer reviews is often used to reinforce and back up already-formed product preferences. Strong positive review texts should therefore be most effective in forming purchase intentions as they strengthen the information seeker’s preferences. Strong negative reviews, on the other hand, pose a threat to his or her quest for consistency. This is likewise the case for reviews that contain both positive and negative information chunks. Inconsistent texts are more prone to cause feelings of uncertainty, risk and dissonance. Due to the psychological mechanisms of the negativity effect, reading negative information leads consumers to postpone their purchase decision until they have obtained sufficient positive information (Tsang & Prendergast, 2009).

**Managerial implications**

In order to enrich the managerial implications of this study, we conducted three short interviews with industry experts. Interviewees A and B have background in both consultancy and
academia. Expert A is the founder and director of an internationally renowned marketing consultancy. Interviewee C holds a managing position at a large European online retailer. Based on these interviews, we derive several implications, which are relevant either (1) during product development or (2) during the after-sales period. Activities relating to the first of these imply a long-term commitment by firms, while implications regarding the second are realised within shorter time spans.

A significant increase in purchase intentions between medium and strong reviews, but no similar effect for negative reviews, was found. Hence, managers must seek ways to provide excellent products and make it easy for consumers to write strong positive reviews. The creation of superior customer value is the cornerstone in this respect (e.g., Gruen, Osmonbekov, & Czaplewski, 2006; Hartline & Jones, 1996; McKee, Simmers, & Licata, 2006; Reichheld, 1996). Yet, it is not only the product/service improvements developed by the companies themselves but more importantly the integration of consumers into co-creation activities and new product developments that will create superior customer value and improve overall success (e.g. running shoes, clothes). Various internet-based tools (e.g., configurators and toolkits) facilitate the engagement and subsequent involvement of consumers. This consumer empowerment will increase people’s experiences of self-determination, efficacy, involvement and even enjoyment. Managers should, therefore, invite consumers to actively participate in the development and discussion of existing or envisaged market offerings (Fueller, Muehlbacher, Matzler, & Jawecki, 2009). This process of co-creation and empowerment will increase loyalty and thus the number of favourable eWoM communications on online networks (Fueller et al., 2009; Kozinets, R.V., de Valck, K., Wojnicki, A.C., & Wilner, S.J.S. (2010)).

After sales, firms should consider the following activities to benefit from the results of this study. First, firms need to identify delighted customers by using tools that are suitable for the
online context. The net promoter score (NPS) developed by Reichheld (2003) is an adequate tool for this purpose. Reichheld suggested measuring the WOM effect on a 0–10 rating scale by asking a simple question: ‘How likely is it that you would recommend our company to a friend or colleague?’ This was in an offline context. The question enables firms to calculate the NPS: the ratio of promoters (9–10) to detractors (0–6) (Reichheld, 2003). As a by-product, the tool helps to identify delighted customers. Since the approach promises simple and timely data that are correlated with the growth and profits of a firm, it can easily be adopted by internet companies. Companies such as Microsoft, General Electric and American Express have already followed the recommendation and integrated ‘the ultimate question’ into their market research activities. Interviewee A confirmed this assessment, but stressed the importance of simultaneously measuring a customer’s willingness to engage in positive eWoM communication.

Second, having identified delighted customers, we suggest actively inviting these persons to write strong positive product reviews. According to Interviewee B, particularly delighted customers need to be stimulated (e.g., by an email following product delivery), as these customers often do not feel compelled to engage in positive eWoM communication, whereas unsatisfied customers ‘naturally’ do. Additionally, strong positive reviews could be placed directly on the package of a product (e.g. on the book cover) or right next to it on online platforms to increase the diffusion and positive effects of eWoM.

Third, the selection of trustworthy online distribution channels becomes of major importance. From a scientific perspective, Hu et al. (2012) underline the issue of fraudulent manipulation of online reviews. In our expert interviews, all of the interviewees were convinced that the number of faked or manipulated eWoM communications is increasing. Thus, online reviews on ‘traditional’ websites such as Amazon or eBay, as well as online communities (e.g., Second Life), social networks (e.g., Facebook) and online blogs, are still of major relevance.
(Fueller et al., 2009; Kozinets et al., 2010) but firms need to critically monitor the management (in terms of technical and security issues) of these websites to ensure that positive and sustainable eWoM communication takes place. Only by doing so will firms be able to identify and address any evidence of negative product attributes or experiences. In the short term this will improve complaint management and after-sales services. In the long term it could even facilitate the creation of superior customer value by firms and initiate valuable co-creation processes.

**Limitations and further research**

The present study does have some limitations: Firstly, our findings are based on an online experiment. Whereas the technique has been criticised for its limited control of situational effects during the experiments, we support the view that the ‘test situation’ should be as realistic as possible. Hence, we think that online experiments are appropriate for a study on online reviews. For a comparison of the results, we suggest that a replication study be carried out in a traditional offline experimental setting. Secondly, we have used three products (books, hotels and running shoes) that are typically bought online. Again, we think this approach is appropriate for an exploratory study, since the research on this topic is still scarce. Finally, we have compared groups exposed to different levels of valence intensity. However, we have not searched for the threshold of strength at which an ‘ineffective’ product review tends to become effective. Although this threshold will certainly be hard to determine, such findings would be of considerable interest to researchers and marketing managers. Testing for this threshold would require a relatively high number of stimuli (reviews) that vary only slightly in valence intensity. Further research on the valence and valence intensity of online reviews should also address the issue of possible personal differences, by including personality variables (e.g. Chung and Rao,
2012). In this regard, we suggest having a closer look at the latent constructs of susceptibility to interpersonal influence (Bearden, Netemeyer, & Teel, 1990) and self-esteem (Rosenberg, 1965). We believe that these concepts would contribute greatly to the understanding of the eWOM effect in the context of online reviews that vary in valence intensity. The explanatory value of other phenomena, already found to be relevant in a broader context of eWOM, such as information overload (Chen et al., 2008) and cognitive personalisation (Xia & Bechwati, 2008), should also be taken into account when investigating the quality of online reviews.
References


### Appendix B: Sample reviews used in Study 2

<table>
<thead>
<tr>
<th>Construct</th>
<th>Source</th>
<th>Cronbach’s Alpha</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How likely is it that you will buy the XXX?</td>
<td>(Sparks &amp; Browning, 2011)</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>How likely is that you will recommend XXX?</td>
<td>(Zeithaml, Berry, &amp; Parasuraman, 1996)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much are you willing to pay for XXX?</td>
<td>(Homburg, Koschate, &amp; Hoyer, 2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Message Credibility</strong></td>
<td>(Cheung et al., 2009)</td>
<td>.825</td>
<td>.826</td>
</tr>
<tr>
<td>I think the online review is credible.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I think the online review is accurate.</td>
<td></td>
<td></td>
<td>.866</td>
</tr>
<tr>
<td>I think the online review is factual.</td>
<td></td>
<td></td>
<td>.855</td>
</tr>
<tr>
<td><strong>Product Category Involvement</strong></td>
<td>(Bloch 1981)</td>
<td>.905</td>
<td>.935</td>
</tr>
<tr>
<td>I have a strong interest in XXX</td>
<td></td>
<td></td>
<td>.929</td>
</tr>
<tr>
<td>I enjoy discussing XXX with my friends</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I actively search for information about XXX</td>
<td></td>
<td></td>
<td>.897</td>
</tr>
<tr>
<td><strong>Perceived Risk</strong></td>
<td>(Forsythe et al., 2006)</td>
<td>.716</td>
<td>.826</td>
</tr>
<tr>
<td>I can’t examine the actual product</td>
<td></td>
<td></td>
<td>.836</td>
</tr>
<tr>
<td>I may not get the right product</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: FL = factor loading. CA = Cronbach’s alpha.
Appendix B: Sample reviews used in Study 2

In this hotel you will find very lovely, clean and well-kept rooms that are newly renovated, include a balcony or terrace, and feature a generous bathroom with hair dryer. The generous all-inclusive services include a very good buffet in the hotel restaurant. The location is also perfect: situated just about 30 km from the airport, the site lies directly on its own clean and well-tended beach. The service staff are friendly, courteous, and even able to speak German, and the rooms are cleaned daily with the beds freshly made. All in all, this is a great hotel and I am very happy to recommend it.

Smiles at the reception area were overall the exception. The staff seemed generally overextended and unorganised. At our arrival we were not offered the room we had booked – the booked direct sea view looked over the roof of the pool taverns and suffered from the respective noise and unpleasant smell. The balcony could only be used when there were offshore winds or else it felt like you were in a fry shack. The meals offered an adequate amount of “edible” selections, but the food was boring and unvaried, drinks were self-serve, the buffet was terribly narrow and cramped, constant bumping and jostling, no decoration…the meals in the main restaurant were a nightmare! The customary fight for the pool loungers around the main pool at 8:30 in the morning! The rooms are old with a variety of defects (odour, furniture condition), and the sanitary facilities seem to be have been there from the time the hotel was established.

The hotel consists of the main building and a number of two-storey bungalows distributed across the large complex. The site is beautifully landscaped and for the most part very nicely distributed. Due to the many pools and pool loungers on the sea it doesn’t feel overcrowded, and one can always find a nice place to relax. Unfortunately, the rooms themselves are old and somewhat run down, but at least they are clean. Although the television didn’t work upon arrival, a very friendly member of staff quickly fixed it. The food wasn’t exhilarating, but ok. All in all, it was a very nice vacation.

The author succeeds in really captivating the reader. With as many details as needed, but without the risk of losing the overview, he narrates with a skillful arc of suspense this multisided, colorful and at the same time terrifying real-life story. Unfortunately, the story can at times seem rather drawn out, but then on the other side is rather predictable. Generally, however, the book is very well done and was a real pleasure to read.