The Good, the Bad, and the Expert: How Consumer Expertise Affects Review Valence Effects on Purchase Intentions in Online Product Reviews

Paul E. Ketelaar  
Behavioural Science Institute / Department of Communication Science, Radboud University, Thomas van Aquinostraat 8, 6525 GD Nijmegen, The Netherlands

Lotte M. Willemsen  
The Amsterdam School of Communication Research, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands

Laura Sleven  
bol.com, Papendorpseweg 100, 3528 BJ, Utrecht, The Netherlands

Peter Kerkhof  
Department of Communication Science / The Network Institute, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam.

This study aims to shed more light on the question whether, and under what circumstances, valence affects consumers’ intention to buy a product after reading an online review. We hypothesize that receiver expertise could possibly moderate (a) the impact of review valence on consumers’ purchase intentions, and (b) the asymmetric effects of positive and negative reviews. To test these hypotheses, we conducted an experiment, exposing participants (n = 470) to reviews varying in valence (i.e., positive, neutral, negative), with purchase-intention as the dependent variable. The results support the moderating role of receiver expertise for both the influence and weight of review valence effects. This explains the inconsistent results for review valence reported in previous studies.

Keywords: Electronic Word of Mouth, Online Reviews, Negativity Effect, Purchase Intention, Consumer Expertise.

doi:10.1111/jcc4.12139

Editorial Record: First manuscript received on August 25, 2014. Revisions received on February 27, 2015, July 20, 2015 and August 20, 2015. Accepted by Miriam J. Metzger on September 2, 2015. Final manuscript received on September 8, 2015.
In the past decade, websites offering customer evaluations of products have grown explosively. These online product evaluations, also referred to as online reviews, typically consist of a recommendation, whether positive or negative. Online reviews have become a popular source for consumers to assess the quality of products because they are perceived as more trustworthy than traditional marketing messages (Bickart & Schindler, 2001). Recent research (Nielsen, 2012) has shown that consumers trust consumer opinions more than any other form of advertising. In fact, a substantial proportion of consumers (24%) consult online reviews before making a purchase decision (Comscore, 2013).

Given the popularity of online reviews, and their importance for both consumers and businesses, a great deal of research has been conducted in the past 10 years to investigate their impact. These studies have mostly focused on valence: i.e., the evaluative tone of the review, which can vary from very negative to very positive (Purnawirawan, Dens, & De Pelsmacker, 2012). For example, in one of the earliest studies, Basuroy, Chatterjee, & Ravid (2003) found online review valence had a significant impact on box office revenues, with positive reviews helping box office revenues and negative reviews hurting them. The study also attested to the commonly held belief that “bad is stronger than good” (Ahluwalia, 2002; East, Hammond, & Lomax, 2008): Negative reviews hurt a movie’s box office performance more than positive reviews improve it.

However, as more studies were conducted, it became apparent that review valence does not always predict purchase behavior. While some studies did find significant effects of review valence, other studies failed to find them. Likewise, while certain studies found negative reviews affecting consumer behavior more strongly than positive reviews, other studies found no asymmetrical effects (for a review, see Wu, 2013; Kimmel & Kitchen, 2014).

These inconsistent findings necessitate a further investigation of the boundary conditions associated with the impact of online reviews (Kimmel & Kitchen, 2014; Wu, 2013). The aim of the present study is therefore to elucidate whether, and under what circumstances, review valence affects the purchase decision-making process. We argue that receiver expertise could possibly moderate (a) the impact of review valence on consumers’ purchase intentions, and (b) the asymmetric effects of positive and negative reviews. To the best of our knowledge, no study so far has explored the moderating role of receiver expertise, here defined as the receivers’ knowledge about the product or product class, derived from prior experience, study, or training (cf. Friedman & Friedman, 1979). Yet, for both theoretical and managerial reasons, it seems imperative to investigate this relatively unexplored link.

From a theoretical perspective, the moderating role of receiver expertise is worth studying. As research in offline contexts has suggested, experts and novices are not equally susceptible to positive and negative messages. According to consumer search and motivated reasoning theory, this is because they differ both in terms of their knowledge structures and their search and processing goals (Lee & Koo, 2012; Bloch, Sherrell, & Ridgway, 1986; Kunda, 1990). While experts are motivated to form accurate judgments and to build “a bank of knowledge,” novices typically lack such extensive knowledge structures, have no interest in building them, and are primarily motivated by the need to arrive at a purchase decision (Bloch et al., 1986). Moreover, category-diagnosticity theory asserts that consumers are more or less inclined to rely on positive and/or negative messages, depending on their processing goals (Ahluwalia, 2002). This study integrates these insights, and compares the effects of positive versus negative online reviews on both expert and novice consumers.

From a managerial perspective it seems important to examine experts’ and novices’ susceptibility to positive and negative information in the context of consumer reviews, as expertise is generally understood to be an important segmentation variable (Chen & Xie, 2008; Xun & Reynolds, 2010). Although experts and novices have different motivations for searching for information, both segments rely on online reviews for product information. Hence, a thorough understanding of how different segments are influenced by online reviews is central for marketers.
Theoretical Foundations and Hypotheses Development

Prior Research on the Impact of Online Review Valence

Review valence is probably the most studied variable in the literature on the effects of online reviews (Cheung & Thadani, 2012). Review valence refers to the tone with which products are being discussed in online reviews, with positively valenced reviews emphasizing a product’s strengths, and negatively valenced reviews emphasizing its weaknesses (Cheung et al., 2009). Review valence can be reflected by numerical rankings (e.g., 5-point star recommendations) or by textual content, and is believed to signal product quality. Hence, review valence functions as a recommendation that can inform consumers’ purchase decisions (Bickart & Schindler, 2001).

Given its signaling function, it is not surprising that most studies of online reviews have focused on valence. Nevertheless, despite the large literature on review valence, the findings on its effects in shaping consumers’ purchase decisions are mixed (Cheung & Thadani, 2012; Cheung, Luo, Sia, & Chen, 2009; Cheung et al., 2009; Lee & Koo, 2012; Lee, Park & Han, 2008; Wu, 2013).

A number of studies have found that review valence significantly influences consumers’ purchase behavior. For example, Utz, Kerkhof and van den Bos (2012) found strong effects of review valence on the evaluation of online retailers. Sen and Lerman (2007) found review valence effects, especially in the case of utilitarian products; and Vermeulen and Seegers (2009) found effects of review valence on the attitude towards hotels. The general conclusion of these and other studies is that online reviews influence purchase behavior in the same direction as the valence of the online reviews; negative information discourages consumers from purchasing particular products or services, while positive information encourages them to make such purchases.

Despite this, other studies have found no effects for review valence (Cheung et al., 2009; Duan, Gu, & Whinston, 2008; Liu, 2006). For instance, in their analysis of user reviews posted on Yahoo! Movies discussion boards, Duan and colleagues (2008) found that review valence did not significantly explain box office revenues (cf. Liu, 2006). Similarly, in a study conducted by Cheung and colleagues (2009) across a variety of products, no significant effects of review valence were found.

Strikingly, most studies that have demonstrated significant effects of review valence have used an experimental design (Lee, Rodgers, & Kim, 2009; Xue & Zhou, 2010), whereas studies that have been unable to find significant effects have used aggregate market-level analysis (e.g., Duan et al., 2008; Liu, 2006). The latter use objective market parameters (e.g., sales and five-star review ratings), which enhances the external validity of these studies. However, unlike experimental studies they cannot control for audience heterogeneity (Chakravarty, Liu, & Mazumdar, 2010). In particular, these studies cannot rule out the possibility that the differences in sales (or the lack thereof) may be due to unobserved differences in the audience that uses online reviews for consumer decision-making. This is an important consideration, given that receivers of online reviews are unlikely to be homogenous. Receivers of online reviews differ in terms of demographics, perceptions, knowledge, experiences, etc. (Chakravarty et al., 2010; Khare, Labrecque, & Asare, 2011; Zhu & Zhang, 2010).

More importantly, message impact is generally understood to result from source, message, and receiver-related factors (Morris & Ogan, 1996). The same message from the same source may produce different effects, depending on the receiver (Chaiken & Eagly, 1976). Although source (e.g., online review vs. other sources of product information) and message factors (e.g., positive vs. negative valence) have received sufficient attention in exploring the effects of online reviews (Cheung & Thadani, 2012), less attention has been devoted to receiver factors. The literature has only recently begun to examine individual-level differences in online review susceptibility, considering dimensions such as internet experience (Zhu & Zhang, 2010), trust disposition (Utz et al., 2012), and need for uniqueness (Khare et al., 2011). These studies provide initial proof that the impact of online reviews is contingent on the
individual experiences, personality traits, and needs of the receiving audience. Hence, the literature calls for more research to examine the issue of how individual-level differences affect consumers' responsiveness to review valence (Chakravarty et al., 2010; Zhang et al., 2010).

**Receiver Expertise as an Individual-Level Moderator of Review Valence Impact**

One individual-level characteristic that could potentially moderate the effects of review valence is receiver expertise, which refers to their knowledge about the product or product class, derived from experience, study, or training (cf. Friedman & Friedman, 1979). Experimental research has shown that receiver expertise affects the processing of consumer information in general (Alba & Hutchinson, 1987), and specifically the processing of online review content such as argument sidedness and factuality (Park & Kim, 2008; Lee & Koo, 2012). For example, Park and Kim (2008) found that expert consumers were more influenced by online reviews focussing on product attributes, whereas less expert consumers were more influenced by reviews focussing on product benefits. Additionally, Lee and Koo (2012) showed that while expert consumers considered objective reviews more credible, both expert consumers and novices considered subjective reviews to be equally credible.

Although these studies support the moderating role of receiver expertise on the effects of review content, none of them specifically focused on the evaluative tone of the reviews (Lee & Koo, 2012; Cheung, Sia, & Kuan, 2012). As a result, whether and how experts and novices differ in their responsiveness to review valence remains unexplored (Cheung & Thadani, 2012). Yet, research on consumer search behavior and information processing provide theoretical arguments that experts and novices would be differently susceptible to the effects of positive and negative reviews.

First, compared to novices, experts have a larger store of knowledge due to prior experience, study or training, and their involvement with the product category (Bloch et al., 1986; Gilly, Graham, Wolfinbarger & Yale, 1998). Experts can be characterized by a continuous interest and enthusiasm for a product category, which may lead them to search for information on an ongoing basis. In contrast, novices do not have an inherent interest in a product category, and are only temporarily involved with knowledge acquisition in order to make a purchase decision (Bloch et al., 1986). Resulting differences in the degree to which both segments accumulate knowledge on the product category may affect consumers' ability to be influenced. Consumers who possess no prior knowledge about a product or product class are likely to rely on any new information they acquire, as this is their only basis for a judgment. Experts, on the other hand, can rely on their own store of knowledge; as a result they will tend to have more confidence in their own judgments and are less likely to be influenced by other people's opinions (Cheung, Xiao, & Liu, 2012). This was convincingly demonstrated by Wood and Lynch (2002), who showed that experts were less likely to change their opinion after encountering new information. Experts were persuaded less than novices, using prior knowledge to generate counterarguments. In contrast, having no or very limited previous knowledge increased the persuasive effect of new information. Similar results were obtained by Gilly and colleagues (1998, Study 2), who examined receivers' susceptibility to word of mouth communication about durable goods, and found that receiver expertise lessened the ability to be influenced.

Another reason to expect differences between experts and novices is based on perceived source similarity, the extent to which individuals perceive a source to be similar to themselves (McCroskey, Richmond, & Daly, 1975). Perceived source similarity has been found to be a key driver of review persuasiveness (Chakravarty et al., 2010). This has been explained by the perceived trustworthiness of similar sources: Individuals simply trust sources more when they perceive them as being similar (Metzger, Flanagan, & Medders, 2010; Smith, Menon, & Sivakumar, 2005). Perceived source similarity may be especially relevant in online contexts where information is exchanged between unknown contacts (for a review, see Walther et al., 2010). As online reviewers are mostly anonymous, receivers often know little about them.
other than that they are (or appear to be) everyday consumers. Since everyday consumers are assumed to have no expert knowledge of the product or category under review (Willemsen, Neijens, Bronner, & de Ridder, 2011), novices may ascribe more similarity to online reviewers than do expert consumers. Such similarity may result in a greater influence of online reviews among novices than expert consumers.

Based on the theoretical differences in terms of knowledge structure and perceived source similarity, we expect that expert and novice consumers will vary in the degree to which they are affected by review valence. Specifically, we expect that novice consumers are more susceptible to online reviews than experts, and as a result will be less likely to purchase a product after reading a negative online review, and will be more likely to buy a product after reading a positive review. Hence, we formulated the following hypotheses:

H1: Positively (vs. negatively) valenced reviews have a positive (vs. negative) effect on purchase intention.

H2: Expertise moderates the effect of online review valence: negative online reviews lead to weaker purchase intentions among novice consumers than expert ones, and positive online reviews lead to stronger purchase intentions among novice consumers than expert ones.

Prior Research on the Asymmetric Effects of Positive Versus Negative Reviews
Mixed findings have not only been reported for the impact of positively and negatively valenced reviews, but also for their relative weight. Although positive reviews outnumber negative reviews (East, Hammond & Wright, 2007), some studies have found that negative reviews outweigh positive ones in terms of their effect: negative reviews have a stronger impact on judgment and behaviors than do positive reviews. This so-called negativity effect has been demonstrated for a wide range of outcomes including trust perceptions, brand evaluations, product choice, and purchase behaviors (Lee, Rodgers, & Kim, 2009; Xue & Zhou, 2010). Moreover, consumers pay more attention to negative reviews than neutral or positive ones (Daugherty, Ward, & Hoffman, 2013), and disseminate negative reviews to more recipients and for a longer period (Hornik, Satchi, Cesareo, & Pastore, 2015).

The tendency of negative information to have a greater impact than positive information is a robust finding in the literature on impression formation. This has been explained by the category-diagnosticity theory, which asserts that people form impressions of others by means of categorization processes (Skowronski & Carlston, 1989, p. 689). During these processes, people categorize others as either good or bad based on diagnosticity judgments: i.e., determining a cue’s usefulness in making distinctions. When forming impressions of people, negative information is considered more useful than neutral or positive information by virtue of its rarity. Negative attributes are less commonly observed in individuals than positive ones (Fiske, 1980), and are more instrumental for classifying individuals as “bad” than positive attributes are for classifying individuals as “good” (Skowronski & Carlston, 1989). Research has shown that these findings also extend to consumers’ evaluation of products, given the predominance of positive product information over negative product information. As most products, whether low-, average-, or high-quality, have positive attributes, negative ones are much rarer. Hence, negative product attributes are considered to be more characteristic of poor-quality products than positive attributes are for high-quality products (Ahluwalia, 2002; Herr, Kardes & Kim, 1991).

Although the negativity effect is well documented, a number of studies on review valence have failed to find a superior effect of negative reviews (Schindler & Bickart, 2012, Ong, 2012; Wu, 2013). Some studies have even reported a positivity effect in which positive reviews weigh more strongly in consumers’ judgments and behaviors than negative ones (East et al., 2008). Interestingly, most of the studies reporting asymmetrical effects in the relative weight of negative versus positive reviews were experimental
(Ahluwalia, 2002; East et al., 2008; Wu, 2013). As mentioned above, the asymmetrical effect of negative reviews only materializes if the negative information in the reviews is considered to be diagnostic and useful. Recent insights reveal that diagnosticity judgments are a subjective assessment dependent on contextual as well as individual factors (Ahluwalia, 2002; Wu, 2013). This leads to the question whether the inconsistent results regarding the weight of negative versus positive reviews could be explained by individual characteristics of the audience. We expect that receiver expertise may not only moderate the presence of review valence effects but also the weight of those effects, due to differences in experts’ and novices’ processing goals.

**Receiver Expertise as an Individual-Level Moderator of Asymmetric Effects**

According to motivated reasoning theory (Kunda, 1990), consumers are generally motivated by two types of goals when processing information: (1) accuracy goals, and (2) precommitment goals. When consumers are motivated by accuracy goals, information processing is driven by the need to arrive at an *accurate* conclusion. When consumers are driven by precommitted goals, information processing is driven by the need to arrive at a *particular* conclusion.

Distinguishing between these goals can help us understand why experimental studies found the most compelling evidence for the presence of negativity effects. Most experimental studies on the effects of review valence have asked participants to form accurate impressions of unknown products and brands, thereby facilitating a relatively impartial treatment of information (East et al., 2007). Interestingly, negative information is considered more diagnostic than positive information when participants engage in impartial processing and are motivated by accuracy concerns (Ahluwalia, 2002). Since expert consumers are more likely to engage in impartial processing than novices due to their inherent interest in building a “bank of knowledge” (Bloch et al., 1986), negative reviews may be weighed more strongly for this group of consumers.

The diagnostic value of negative information is expected to be lower for novices than for experts, as these consumers likely do not have such strong accuracy motives when processing information compared to experts. This implies that negative information may be less useful for them than positive information. Ahluwalia (2002) showed that the negativity effect is attenuated, or even reversed into a positivity effect, if the perceived diagnosticity of negative information is lowered. Hence, the negativity effect is less likely to occur for novices than for experts.

Based on this reasoning, we expect that the negativity effect depends on consumers’ level of expertise. Drawing on category-diagnosticity theory and motivated-reasoning theory, the negativity effect is expected to be more pronounced among high expertise participants. Hence, we hypothesize:

H3: Negatively valenced reviews affect purchase intention to a larger extent than positive ones.

H4: Expertise moderates the asymmetrical effects of review valence: whereas negative reviews affect purchase intention to a larger extent than positive reviews among expert consumers, no such negativity effect appears among novices.

**Method**

**Design and Participants**

We conducted an experimental study to test the combined effects of consumer expertise and review valence. Hypotheses were tested using a 3 x 2 mixed model factorial design, with valence (positive, negative, neutral) as a within-subjects factor, and expertise as a between-subjects factor. We chose cameras
as the experimental product as cameras are frequently purchased online, and as this product category has the most reviews after mobile phones. Respondents were exposed to the product descriptions of six cameras: three test products that were used for hypothesis testing, and three filler products. Filler products were included to camouflage the purpose of the study and to minimize order effects.

Review valence was manipulated and expertise was measured. In order to manipulate review valence, each of the test product descriptions was accompanied by either three positive, three neutral, or three negative reviews. Each filler product included three random reviews, exposing respondents to three reviews for each product, instead of a single review for each product, mimics reality better (Purnawirawan et al., 2012) and makes the results less dependent on the specific characteristics of the product and thus more robust. Also, we added a control group (no review) to establish negativity effects by comparing purchase intention after reading a negative or positive review with the mean purchase intention of the same product in a no-review condition.

In total, 470 people participated in the experiment, 115 of them assigned to the control group. Most respondents (49%) were younger than 25, 46.9% were male, 53.1% female. The majority of the respondents were highly educated: 31.9% of the respondents attended higher vocational education, and 38.5% of the respondents were attending university. Information about sample recruitment is provided below.

Stimulus Materials
The stimulus materials consisted of online reviews from the Dutch comparison site vergelijk.nl, that were adjusted for the purpose of this study. Specifically, the valence of the reviews was systematically manipulated according to their ratings and textual content. The average rating in a positive review was always higher than 8 (on a scale of 1 to 10), a negative review never scored higher than an average of three, and a neutral review scored an average of around five. The textual content of the reviews included adjectives that were replaced with positive or negative words to create positive or negative reviews, or neutral words to create neutral reviews.

The reviews always contained information about camera features, such as the number of megapixels. Following Park, Lee, and Han (2007) the product descriptions did not mention the brand of the camera or its price, but focused on its technical qualities. The name of the website on which the reviews were shown—i.e., vergelijk.nl—was removed as well as any other recognizable feature that could be traced back to this specific website. The camera features were randomized and photo connoisseurs were consulted in order to ensure that the camera descriptions were believable (See Figure 1 for an example). A pretest among 10 people indicated that the reviews were indeed perceived as reliable, realistic, and useful, and that the positive valence was experienced as positive ($M = 9.13, SE = 0.69$), the negative valence as negative ($M = 1.60, SE = 0.83$), and the neutral valence as neutral ($M = 4.73, SE = 1.24$). Another test was conducted to check whether the reviews were believed to derive from everyday consumers with no expert knowledge. To measure perceived expertise of the reviewers, respondents in a post test who did not participate in the main study but belonged to the same population ($n = 33$), answered three questions: What applies to the writers of these reviews: (1) “They use standard camera settings” / “They switch between standard camera settings and manual settings” / “They always use manual camera settings” (2) What applies to the writers: “They have never taken a photography course” / “They learned some things about photography from the internet, an acquaintance or some brief lessons” / “They took a photography course,” and finally (3) What applies to these writers: “For them photography is not a hobby or their profession” / “Photography is a hobby” / “Photography is their profession” (cf. definition Friedman & Friedman, 1979). The item scores were averaged to obtain a measure of perceived expertise, with higher scores denoting perceptions of higher expert knowledge ($\alpha = .59$). According to the results, respondents attributed no expert knowledge to the writers of reviews, given the relatively low score on
<table>
<thead>
<tr>
<th>Product description:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Megapixels</td>
</tr>
<tr>
<td>Lens</td>
</tr>
<tr>
<td>Optical zoom</td>
</tr>
<tr>
<td>Width LCD-screen</td>
</tr>
<tr>
<td>Autofocus</td>
</tr>
<tr>
<td>Image stabilizer</td>
</tr>
<tr>
<td>Extra</td>
</tr>
</tbody>
</table>

**Review 1**
I am very satisfied with the camera. I knew the previous model but this model is better than the previous camera on all specs. It weighs less, is compact and therefore handy to take with you everywhere. It makes beautiful pictures, a real top-camera!

<table>
<thead>
<tr>
<th>Picture quality</th>
<th>9</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possibilities</td>
<td>8</td>
<td>grade</td>
</tr>
<tr>
<td>Ease of use</td>
<td>8</td>
<td>8.3</td>
</tr>
</tbody>
</table>

**Review 2**
Marvelous camera! It provides a good grip, and is easy to use. The first pics already look awesome. The camera has many features, which you get to know in an instant. Great use of nice material.

<table>
<thead>
<tr>
<th>Picture quality</th>
<th>10</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possibilities</td>
<td>9</td>
<td>grade</td>
</tr>
<tr>
<td>Ease of use</td>
<td>8.5</td>
<td>9.2</td>
</tr>
</tbody>
</table>

**Review 3**
Bright screen, good menus, fast with AF and GPS, good battery life. The screen on top is great too. It has a sealable viewfinder for long shutter speeds. I am very satisfied with this purchase.

<table>
<thead>
<tr>
<th>Picture quality</th>
<th>9</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possibilities</td>
<td>9</td>
<td>grade</td>
</tr>
<tr>
<td>Ease of use</td>
<td>8</td>
<td>8.7</td>
</tr>
</tbody>
</table>

**Figure 1** Product description of a camera with three positive reviews.

the scale ($M = 1.73, SD = 0.38$). This was confirmed by an one-sample t-test, which showed that the mean score was significantly below the middle category of the scale, $t(32) = -4.157, p < .001$.

**Procedure**
Respondents were approached in various ways to ensure a population that varied in their camera/photography expertise. A message was posted on three photo forums and we used snowball sampling on Facebook and Twitter to recruit participants. Respondents who agreed to participate in the experiment read the following scenario: “Imagine that you want to buy a new DSLR camera, and that money is not an issue. Please have a look at the following product descriptions and assess for each on a 5-point scale how inclined you would be to buy the camera.” Respondents were then randomly assigned to one of four groups: three experimental groups and one control group. The experimental groups were exposed to three focal product descriptions and three filler product descriptions. The specific pairing of a camera with review valence differed for each experimental group, resulting in a design in which in the different conditions the same cameras were combined with either three positive, three neutral, and three negative product reviews, thus enabling us to rule out both camera and camera x review valence effects. The control group was exposed to the same six product descriptions as the experimental groups, but without any reviews: participants only saw the product description and based their purchase intention exclusively on this. All respondents subsequently rated their purchase intention for each camera, and answered questions about their knowledge of cameras and photography.
Measurement
To measure purchase intention, respondents rated their intention to purchase the reviewed product on a 5-point scale (1 = not inclined to buy, 5 = inclined to buy). To measure consumer expertise, respondents were asked to answer three questions on a 3-point scale: (1) What applies to you: “I use standard camera settings” / “I switch between standard camera settings and manual settings” / “I always use manual camera settings,” (2) What applies to you: “I have never taken a photography course” / “I learned some things about photography from the internet, an acquaintance or some brief lessons” / “I took a photography course,” and finally (3) What applies to you: “For me photography is not a hobby or my profession” / “Photography is a hobby” / “Photography is my profession” (cf. definition Friedman & Friedman, 1979). The item scores were averaged to obtain a measure of expertise ($\alpha = .77, M = 1.70, SD = 0.64$).

Results
Table 1 shows the means and standard deviations of purchase intention for the negative, neutral, and positive review valence conditions and the control group. In order to test our hypotheses, both the control group, the scores of the continuous moderator variable, the three experimental groups, and the three different cameras that were paired with either positive, neutral or negative reviews need to be included in the analysis. Although a within-subjects analysis of variance is the most commonly used way of analyzing a (partly) within-subjects design, this method is less suited when combined with a continuous moderator, and when the experimental setup is not fully balanced (in our design the control group was shown cameras without reviews).

Since our hypotheses require the scores in the three experimental groups to be contrasted with the scores in the control group (H3, H4), a multilevel regression analysis is most suited to analyze the data. In such an analysis, the three scores for the different cameras need to be treated as separate cases. Since the purchase intention scores for the three cameras are nested within individuals, we first tested how much of the variance in purchase intention resides in individuals. An empty variance components model multilevel analysis revealed that due to the strong effect of review valence on purchase intention the within-person variance is not distinguishable from zero. Given the lack of within-person variance, we included only fixed effects in our analysis, which yields results similar to ordinary linear regression analysis. In this analysis, positive, neutral, and negative review valence were entered as (centered) dummies and contrasted with the base rate purchase intention scores of the control (no review) group. Also, centered dummy variables were created to control for the effects of the three cameras (using the scores for the second camera B as base rate). Interaction terms were created to test our expertise x review valence hypotheses, and to control for other possible interactions (camera x review valence; camera x expertise; camera x review valence x expertise). In addition, we controlled for age and sex differences. The results

<table>
<thead>
<tr>
<th>Valence</th>
<th>Camera A</th>
<th>Camera B</th>
<th>Camera C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (n = 112)</td>
<td>1.28 (0.57)</td>
<td>1.22 (0.64)</td>
<td>1.42 (0.89)</td>
<td>1.31 (0.72)</td>
</tr>
<tr>
<td>Neutral (n = 100)</td>
<td>1.91 (0.98)</td>
<td>2.30 (0.97)</td>
<td>2.21 (0.87)</td>
<td>2.14 (0.95)</td>
</tr>
<tr>
<td>Positive (n = 114)</td>
<td>4.08 (1.05)</td>
<td>3.71 (1.30)</td>
<td>3.92 (1.21)</td>
<td>3.90 (1.20)</td>
</tr>
<tr>
<td>Control (n = 115)</td>
<td>2.96 (1.28)</td>
<td>2.69 (1.25)</td>
<td>2.77 (1.16)</td>
<td>2.81 (1.07)</td>
</tr>
</tbody>
</table>
Table 2  Regression analysis of Purchase Intention (n = 1323)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.500</td>
<td>0.176</td>
<td>0.000</td>
<td>2.155</td>
<td>2.846</td>
</tr>
<tr>
<td>Age</td>
<td>-0.151</td>
<td>0.048</td>
<td>0.002</td>
<td>-0.245</td>
<td>-0.057</td>
</tr>
<tr>
<td>Seks (M/F)</td>
<td>0.200</td>
<td>0.076</td>
<td>0.009</td>
<td>0.051</td>
<td>0.350</td>
</tr>
<tr>
<td>Camera A</td>
<td>0.015</td>
<td>0.068</td>
<td>0.821</td>
<td>-0.119</td>
<td>0.149</td>
</tr>
<tr>
<td>Camera C</td>
<td>0.063</td>
<td>0.068</td>
<td>0.353</td>
<td>-0.071</td>
<td>0.197</td>
</tr>
<tr>
<td>Positive review</td>
<td>1.106</td>
<td>0.078</td>
<td>0.000</td>
<td>0.953</td>
<td>1.259</td>
</tr>
<tr>
<td>Negative review</td>
<td>-1.503</td>
<td>0.078</td>
<td>0.000</td>
<td>-1.657</td>
<td>-1.350</td>
</tr>
<tr>
<td>Neutral review</td>
<td>-0.665</td>
<td>0.078</td>
<td>0.000</td>
<td>-0.818</td>
<td>-0.511</td>
</tr>
<tr>
<td>Expertise</td>
<td>-0.027</td>
<td>0.054</td>
<td>0.621</td>
<td>-0.132</td>
<td>0.079</td>
</tr>
<tr>
<td>Expertise * Positive review</td>
<td>-0.604</td>
<td>0.126</td>
<td>0.000</td>
<td>-0.852</td>
<td>-0.357</td>
</tr>
<tr>
<td>Expertise * Negative review</td>
<td>0.326</td>
<td>0.126</td>
<td>0.010</td>
<td>0.078</td>
<td>0.573</td>
</tr>
<tr>
<td>Expertise * Neutral review</td>
<td>-0.060</td>
<td>0.126</td>
<td>0.632</td>
<td>-0.308</td>
<td>0.187</td>
</tr>
<tr>
<td>Camera A * Positive review</td>
<td>-0.093</td>
<td>0.190</td>
<td>0.624</td>
<td>-0.466</td>
<td>0.280</td>
</tr>
<tr>
<td>Camera A * Negative review</td>
<td>-0.293</td>
<td>0.193</td>
<td>0.129</td>
<td>-0.671</td>
<td>0.085</td>
</tr>
<tr>
<td>Camera A * Neutral review</td>
<td>-0.639</td>
<td>0.191</td>
<td>0.001</td>
<td>-1.014</td>
<td>-0.264</td>
</tr>
<tr>
<td>Camera A * Expertise</td>
<td>-0.019</td>
<td>0.109</td>
<td>0.861</td>
<td>-0.232</td>
<td>0.194</td>
</tr>
<tr>
<td>Camera C * Positive review</td>
<td>0.001</td>
<td>0.191</td>
<td>0.996</td>
<td>-0.374</td>
<td>0.376</td>
</tr>
<tr>
<td>Camera C * Negative review</td>
<td>-0.009</td>
<td>0.190</td>
<td>0.963</td>
<td>-0.382</td>
<td>0.364</td>
</tr>
<tr>
<td>Camera C * Neutral review</td>
<td>-0.196</td>
<td>0.193</td>
<td>0.310</td>
<td>-0.574</td>
<td>0.183</td>
</tr>
<tr>
<td>Camera C * Expertise</td>
<td>0.093</td>
<td>0.109</td>
<td>0.392</td>
<td>-0.120</td>
<td>0.306</td>
</tr>
<tr>
<td>Camera A * Positive review * Expertise</td>
<td>0.481</td>
<td>0.311</td>
<td>0.122</td>
<td>-0.129</td>
<td>1.092</td>
</tr>
<tr>
<td>Camera A * Negative review * Expertise</td>
<td>0.065</td>
<td>0.316</td>
<td>0.838</td>
<td>-0.555</td>
<td>0.684</td>
</tr>
<tr>
<td>Camera A * Neutral review * Expertise</td>
<td>0.520</td>
<td>0.298</td>
<td>0.082</td>
<td>-0.066</td>
<td>1.105</td>
</tr>
<tr>
<td>Camera C * Positive review * Expertise</td>
<td>0.261</td>
<td>0.298</td>
<td>0.382</td>
<td>-0.324</td>
<td>0.847</td>
</tr>
<tr>
<td>Camera C * Negative review * Expertise</td>
<td>0.081</td>
<td>0.311</td>
<td>0.793</td>
<td>-0.529</td>
<td>0.692</td>
</tr>
<tr>
<td>Camera C * Neutral review * Expertise</td>
<td>0.309</td>
<td>0.316</td>
<td>0.328</td>
<td>-0.310</td>
<td>0.928</td>
</tr>
</tbody>
</table>

(see Table 2) show no main effects of the specific camera, indicating that the three cameras were equally attractive to the respondents. Also, only 1 out of 14 two- or three-way interactions with camera, expertise and review valence was significant, indicating that the specific camera and the combinations of camera and review played no role in the outcomes. The analysis also shows that purchase intention was higher among younger and female respondents.

Positive and negative review valence both strongly affected purchase intention in the predicted direction (resp. B = 1.106, S.E. = .078; and B = - 1.503, S.E. = .078), which confirms H1. Unexpectedly, a neutral review had a negative effect when compared to base rate of no review (B = -0.665, S.E. = .078), indicating that the neutral review may not have been perceived as neutral.

H2 stated that negative online reviews will lead to weaker purchase intentions among novice consumers than expert ones, and positive online reviews to stronger purchase intentions. This hypothesis was tested through the (positive and negative) review valence x expertise interaction. Both interaction terms are significant and in the predicted direction. To illustrate the results, we conducted a median split
on expertise ($M_{\text{low expertise}} = 1.08; M_{\text{high expertise}} = 2.20$) and plotted the means of purchase intention for positive and negative reviews and the no review condition for both levels of expertise (Figure 2). The pattern of means reveals that indeed for both negative and positive reviews, the effects of review valence are stronger among respondents with low levels of expertise.

Our third hypothesis H3 stated that negative reviews affect purchase intention more than positive ones (the negativity effect). We compared the slopes for positive and negative reviews (vs. no review) and tested whether the $B$ values are larger for negative reviews (H3) by calculating a $z$-score, $z = \frac{(\text{abs}(B_1) - B_2)}{\sqrt{(SEb_1^2 + SEb_2^2)}}$. This results in a score of $z = 3.60, p < .0001$, indicating that the effect of a negative review does indeed exceed the effect of a positive review.

The negativity effect was expected to be more pronounced among high expertise participants, meaning that negative reviews will affect purchase intention to a larger extent than positive ones among expert consumers, but not among novices (H4). To determine this, we must contrast the slopes of both positive and negative reviews among low and high expertise respondents. As noted before, review valence, both positive and negative, exerts a stronger effect in general among novice consumers. Figure 2 showed that the absolute values of the effects for negative and positive reviews are indeed similar for low levels of expertise, whereas for high levels of expertise the effect of negative reviews is steeper than the effect of positive reviews. We compared the slopes of the regression lines for high ($+1$ SD) and low ($-1$ SD) expertise respondents and again computed $z$-scores to test the difference. The results in Table 3 show that the negativity effect is only found among expert respondents, with the $B$ for negative reviews more than three times the size of the $B$ for positive reviews ($z = 12.01, p < .001$). Among novice respondents, there is a smaller but significant positivity effect ($z = 4.817, p < .001$), indicating a somewhat stronger effect of positive (vs. negative) reviews. Although the latter finding is unexpected, the findings confirm H4 because the negativity effect is only found among expert respondents.

**Discussion**

The purpose of the current study was to shed more light on how consumer expertise moderates online review and review valence effects on the purchase decision-making process. Our experimental study
Specifically aimed to assess the effects of online review valence on purchase intention and the moderating role of consumer expertise.

The results show that online reviews affect purchase intention in the same direction as their valence: Positive reviews have a positive effect on purchase intention, whereas negative reviews have a negative effect, as compared to both a control condition and a neutral review (H1). This is in line with the findings of earlier studies that have adopted an experimental design (e.g., Sen & Lerman, 2007; Utz et al., 2012; Vermeulen and Seegers, 2009), but contradicts the outcomes of field studies that used real-world data to document the effects of online reviews (Cheung et al., 2009; Duan et al., 2008; Liu, 2006). A potential explanation, as pointed out and tested by the current study, is that field studies are not able to control for individual differences in consumer expertise that may influence the effects of online review valence. The results do indeed seem to support this contention. As convincingly demonstrated, consumer expertise moderates the effects of online review valence. The effects of positive and negative reviews on purchase intention were generally stronger for novices (H2). Compared to experts, novices judged the reviewed product more negatively after reading a negative review, and more positively after reading a neutral or a positive review.

Additionally, we find that negative valence is weighed more heavily in terms of purchase intention than positive valence (H3). Such asymmetrical effects have been reported in other experimental studies, but not or less convincingly in field studies (Lee & Koo, 2012; Wu, 2013). However, it is essential to take individual differences into account, and we found that the relative influence of positive and negative reviews is mitigated by consumer expertise: Whereas novice consumers actually showed a slight positivity bias, expert consumers had a negativity bias, being less affected by positive reviews than by negative reviews (H4). This negativity bias among experts can be explained by the idea that they are intrinsically more interested in the product category than novices and, following Bloch et al. (1986) and Ahluwalia (2002), engage in information banking, which results in more impartial processing and higher accuracy concerns.

The unexpected positivity effect that we found among novice respondents corroborates the literature about diagnosticity of evaluative information (Ahluwalia, 2002). Because experts more than novices are driven by accuracy goals, they likely pay more attention to negative reviews. In fact, in some cases novices may be driven more than experts by precommitment goals (Bloch et al., 1986), which is the desire to purchase a particular product. In precommitment, consumers are found to process information in a goal-supporting manner (Khare et al., 2011; Kunda 1990). This implies that negative information will be less informative for them than positive information, as negative information about a product is not supportive of a purchase decision. Indeed, Ahluwalia (2002) showed that the negativity effect is attenuated or even reversed into a positivity effect if the perceived diagnosticity of negative information is lowered. If novices are driven by these goals whereas experts are driven by accuracy goals, this could explain why novices are influenced more by positively framed reviews. Thus, when controlling for audience heterogeneity, this study confirms the presence of the negativity bias: Receivers tend to weigh negative information more strongly in their decision to purchase products than positive information (Skowronski & Carlston, 1989).

### Table 3 Regression slopes of negative and positive (vs. no) reviews for low and high expertise

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$B_{\text{negative review}}$</th>
<th>$B_{\text{positive review}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low expertise (−1 SD)</td>
<td>2.527</td>
<td>−1.178</td>
<td>1.710</td>
</tr>
<tr>
<td>High expertise (+1 SD)</td>
<td>2.474</td>
<td>−1.829</td>
<td>0.502</td>
</tr>
</tbody>
</table>
Theoretical and Managerial Implications

These findings contribute to the literature in two important ways. First, we address the call for more research on individual-level variables as determinants of review valence effects (Chakravarty et al., 2010; Zhang, Cracium, & Shin, 2010). There is a paucity of research on the individual-level differences in consumers’ susceptibility to online reviews. This can be explained by the predominant use of market-level data to examine the link between valence and consumer behavior (Chakravarty et al., 2010). By testing expertise as a moderator of review valence effects, this study contributes to the literature, considering the effects of online reviews from the perspective of the receiving audience. The results demonstrate that not every consumer is equally affected by online review valence. Depending on consumers’ knowledge of the reviewed product class, consumers may be more or less affected by positive or negative online reviews. As such, this study offers important insights into how various consumer segments respond to online reviews.

Second, our findings inform the ongoing debate on the effects of online review valence. Although many studies have addressed this topic, as noted, the results of these efforts have been mixed. This study demonstrates that consumer expertise affects both the extent of the influence of positive and negative reviews, and their relative weight. By testing whether levels of expertise form potential boundary conditions for the effects of review valence, we attempt to reconcile the inconsistent findings found previously. In doing so, this study moves beyond earlier studies that aimed to establish the impact of online reviews and, instead, aims to elucidate the circumstances in which online reviews are likely to influence consumers.

These results also have important practical implications. Perhaps most important is that the conventional wisdom “bad is stronger than good,” which still guides many marketing decisions (East et al., 2008; Wu, 2013), does apply under some circumstances, but certainly not all. Consumers’ knowledge of the reviewed product seems to be relevant: Experts, but not novices, weigh negative reviews more strongly than positive ones. Marketers may take comfort in this finding, as experts generally represent a small proportion of the total consumer population. However, this does not mean that marketers should ignore negative reviews. Instead, marketers can use the insights of this study to manage the circulation and effects of online reviews for specific audiences.

If experts are more likely to be affected by negative reviews than by positive ones, marketers can introduce reward-based referral programs on platforms that are often visited by expert consumers. Referral programs encourage satisfied customers to spread positive word-of-mouth, for example in the form of positive reviews, which may offset the circulation of negative reviews. Marketers can also attempt to offset the effects of negative reviews by publicly responding to negative reviews, a practice referred to as webcare (Van Noort & Willemsen, 2012). If marketers are able to resolve the issues described in negative reviews, they may mitigate the effects amongst expert consumers (Van Noort, Willemsen, Kerkhof, & Verhoeven, 2014). This is important since experts are typically more involved and may be more inclined to write online reviews than novices.

Limitations and Suggestions for Further Research

This study has limitations, which suggest opportunities for future research. The first concerns the ecological validity of this study. Our conclusions are based on a single study in a single product category, implying that generalizing about the effects of consumer expertise both within this product category and across product categories should be done with care. The reviews that were used in this study focused on utilitarian products and more specifically on their objective product attributes (e.g., the brightness of the pictures, dynamic reach, etc.). The findings confirm earlier work by Sen and Lerman (2007) that showed negativity effects for utilitarian products. Future research should establish whether consumer expertise
also moderates the effects of valence in the case of hedonic products or more subjective reviews. There is reason to expect that this may not be the case: prior knowledge of the reviewed product class is less important when product selection is a matter of taste (Lee & Koo, 2012). Hence, expertise may play a less important role as a moderator. Furthermore, future research might examine how experts and novices evaluate mixed reviews that contain both positive and negative information about a single product.

Another limitation regarding the ecological validity concerns the number of reviews that were used in the present study. Each product description was accompanied by three reviews that were positively, neutrally, or negatively valenced, while review volume is much larger and review valence more varied in real life. In fact, Khare et al. (2011) studied review volume as a factor and operationalized high volume with 3,470 reviews and low volume with 62 reviews. Among other things, they found that high volume of posted reviews accentuates perceptions of positivity and negativity of electronic word-of-mouth (eWOM) information. Therefore, future studies could test these relations with higher review volume.

A further shortcoming is that we have only assumed – but not actually tested – perceived diagnosticity as an underlying process that leads to different effects based on consumers’ level of expertise. Although various studies show that perceived helpfulness, as a proxy of diagnosticity, does explain the effects of online review valence (Sen & Lerman, 2007), follow-up research is needed to test whether helpfulness also accounts for the different impact of positive and negative online reviews for experts and novices. The same argument holds for perceived similarity as an assumed underlying mechanism. The present study suggests that novices ascribe more similarity to sources of online reviews than expert consumers, since online reviews are generally believed to be written by everyday consumers like themselves with no expert knowledge, which was also confirmed in our post test. Since consumers are influenced more by similar others (Brown, Broderick & Lee, 2007), ascribed similarity may lead to a greater influence of review valence among novices than expert consumers. To demonstrate such a mechanism, future research needs to test the potential interaction between review valence, receiver expertise and sender expertise, and perceived similarity as a mediator of any effects that are found.

Finally, we measured expert knowledge in a rather indirect way. This indirect measure assumes that experts are more likely than novices to use manual settings, have taken a photography course, and that photography is their profession compared to novices (cf. Friedman & Friedman, 1979). This measure was chosen over self-reported measures of perceived expertise, since people are not always accurate judges of their knowledge. Research shows that people judge their knowledge too favorably, which leads to overestimation of perceived expertise (for a review, see Atir, Rosenzweig, & Dunning, 2015). Although this measure of consumer expertise appears to be useful in distinguishing people with higher versus lower expertise (see results), future research might employ a more direct measure for expertise, by assessing actual knowledge.

Despite these limitations, the study provides initial proof that individual-level differences in receiver expertise affect consumers’ susceptibility to online review valence. As such, this study complements two streams of research on online reviews: (1) *market-level analyses* on the effects of *review valence* that have so far treated consumers as a homogeneous audience; and (2) *individual-level analyses* that have considered the influence of individual differences such as *receiver expertise* on responses to reviews, but never in relation to review valence. This study unravels the effects of differences in consumer expertise on the response to positive, neutral, and negative online reviews. Clearly, receiver expertise is an important factor that needs to be considered when predicting the effects of online reviews.

**References**


About the Authors

Paul E. Ketelaar, Ph.D., is a Senior Assistant Professor of Communication and Influence in the Behavioural Science Institute (BSI), Department of Communication Science, at the Radboud University, Nijmegen, The Netherlands, where he heads the Tailored Communication Project. His research interests focus on cognitive and social issues for the design and use of interactive communication technologies and e-commerce. His research examines tailored communication in virtual reality settings, online strategic communication through social media, and message design of online content. 

*Address:* Behavioural Science Institute / Department of Communication Science, Radboud University,
Lotte M. Willemsen, Ph.D., is professor Crossmedia Business in the department of Communication & Journalism at the HU University of Applied Sciences, Utrecht. Her research focuses on the empowering role of social media from the perspectives of both consumers and organizations. **Address:** Centre for Communication & Journalism, HU University of Applied Sciences, Padualaan 99, 3584 CH, Utrecht, The Netherlands **Email:** lotte.willemsen@hu.nl.

Laura Sleven has completed her Master’s Degree in Communication Science at the Radboud University and is currently working for bol.com where she continues to explore her interest in online marketing. **Address:** bol.com / Papendorpseweg 100, 3528 BJ Utrecht, The Netherlands. **Email:** laura_sleven@hotmail.com.

Peter Kerkhof, Ph.D., currently holds the chair in social media at the Department of Communication Science, VU University Amsterdam, The Netherlands. His recent research interests focus on the uses of social media by consumers and companies, and on the conversations that they engage in on social media. **Address:** Department of Communication Science / The Network Institute, Vrije Universiteit Amsterdam. De Boelelaan 1105, 1081 HV Amsterdam. **Email:** p.kerkhof@vu.nl.