Is a Picture Always Worth a Thousand Words? Attention to Structural Elements of Ewom For Consumer Brands Within Social Media

Ernest Hoffman, University of Akron, USA
Terry Daugherty, University of Akron, USA

The growing influence of social media on consumer judgments makes it important to know what captures consumer attention. We study attention using eye-tracking in the context of social media and consumer-generated Word-of-Mouth. Our results suggest that consumer attention within social media is significantly influenced by brand utility and message valence.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/1014817/volumes/v41/NA-41

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/
Is a Picture Always Worth a Thousand Words? Attention to Structural Elements of eWOM for Consumer Brands within Social Media
Ernest Hoffman, University of Akron, USA
Terry Daugherty, University of Akron, USA

ABSTRACT
Consumer attention has long been known to influence evaluations of, and responses to, advertising stimuli in meaningful ways (e.g., Krugman, 1971; Morrison & Dinoff, 1972). Indeed, behavioral eye-tracking studies have been utilized for quite some time to link attention to outcomes of interest such as product recognition (Kroebner-Riel & Barton, 1980), recall (Krugman, Fox, Fletcher, Fischer, & Rojas, 1994; Rosbergen, Pieters, & Wedel, 1997), and product sales (Treistman & Gregg, 1979). Although tremendous insights have been gained from this work, the context of consumer attention has changed dramatically in at least two ways that necessitate novel exploration.

First, the emergence of social media (e.g., Facebook, YouTube, Pinterest) has created a context where massive amounts of image and text-based messages compete for consumer attention like never before. According to a recent report (Bennett, 2012), consumers share more than 600,000 pieces of content and create over 25,000 social media posts every sixty seconds. In such an over-saturated environment, text and image elements must compete to get noticed at all. This is significant because in order for message elements to have an impact on consumers (i.e., make a difference), they must first be noticed. Yet, attention is typically studied in offline contexts, such as print (e.g., Lohse, 1997; Pieters & Wedel, 2004), where the number of visual and textual stimuli competing to get noticed is substantially more limited.

Second, the growing influence of consumer-generated electronic word-of-mouth (eWOM) has given consumers an unprecedented stake in product advertising (Chu & Kim, 2011; Rieger, 2007; Zhang & Daugherty, 2009; Henning-Thurau, Gwinner, Walsh, & Gremler, 2004). In general, word-of-mouth (WOM) is a critical component of marketing (Brown & Reingen, 1987; Buttle, 1998) with user-generated reviews perceived as more credible and less biased than company-generated advertising (Dellarocas, 2003; Ha, 2006; Keller, 2007; Phelps, Lewis, Mobillo, Perry, & Ramman, 2004). In fact, new customer acquisition (Doyle, 1998), increased sales (Chevalier & Mayzlin, 2006), and product-use decisions (Foster & Rosenweig, 1995; Godes & Mayzlin, 2004) have all been directly linked to WOM communication. Marketers now have the opportunity to initiate and influence WOM unlike ever before on a large scale quickly and efficiently. Subsequently, interest has grown regarding the conditions that dictate whether or not consumers pay attention to WOM (Daugherty and Hoffman in press).

One area of consumer attention that is broadly affected by these changes pertains to structural differences in message presentation, including the strategic use of image versus text-based elements. In traditional offline contexts, where advertising is driven by corporate marketing, image-based elements have been shown to capture attention more than text (e.g., Pieters & Wedel, 2004). However, this may not always be the case in social media contexts where product reviews are consumer-driven, creating an element of uncertainty regarding message valence. In particular, the emergence of negative word-of-mouth (NWOM) introduces the possibility of product reviews being harmful as opposed to promotional.

To further explore these dynamics, we first review the existing literature regarding consumer attention to structural elements of eWOM. We then provide a theoretical rationale for presuming that a more complex relationship unfolds in a social media environment where product information is consumer-generated. A series of testable hypotheses are presented, and an eye-tracking study is undertaken to better understand how visual elements, message valence and brand type interact to influence consumer attention within social media. Implications and managerial recommendations are then discussed on the basis of our findings.

CONCEPTUAL FRAMEWORK
Two prevailing paradigms are used to study how consumers attend to structural elements of advertising. One paradigm is concerned with optimizing the effectiveness of ad component presentation. Factors such as size or location are manipulated to maximize the impact of text or image on attention. For instance, Kroebner-Riel and Barton (1980) manipulated the positioning of text-based elements and found that locating text at the top of an ad leads to better information acquisition. A similar approach was taken by Garcia, Ponsada, and Esteban (2000), who found that image positioning matters, but only when pre-existing product involvement is low. Other studies have considered text and image elements together, manipulating the use of color and graphics (Lohse, 1997) or size and location (Dreze & Huss Herr, 2003) to determine the type of presentation that is most likely to capture attention.

A second approach has sought to compare the effectiveness of images and text in capturing consumer attention. Pieters and Wedel (2004), for instance, found that attention to pictorial elements of ads is superior to text-based elements regardless of manipulations in size. In an earlier study, these authors determined that attention fixation to brand and pictorial elements of ads led to better memory of the brand, whereas text fixations did not. To date, a long line of research has established the relative superiority of images in evoking attention (e.g., Rossiter, 1981; Singh, Lessig, Kim, Gupta, & Hocutt, 2000).

Social media has changed the manner in which consumers are exposed to product-related information. Specifically, it is now possible for products to be linked to multiple different images and textual reviews by eWOM generators, rather than a select handful of marketer-driven formats. To our knowledge however, a comparative integrated paradigm has yet to be adopted within the unique context of consumer-generated content. One major difference in this venue is the greater propensity for consumer-generated message elements to vary in the extent to which they endorse a brand. Nevertheless, without accounting for more complex dynamics (as we do below) we suspect that image-based elements will continue to appear as if they capture more attention amongst social media users than text-based elements.

Qualifying Attention to Structural Elements: Brand Luxury
Brands are frequently distinguished in terms of having either utilitarian or hedonic value (Sen & Lerman, 2007). Utilitarian brands are evaluated by consumers in terms of their usefulness or practicality (Strahilevitz & Meyers, 1998) while hedonic brands are judged in terms of the extent to which they satisfy consumer wants (Hirschman & Holbrook, 1982). Research has shown that assessments of utilitarian brands are cognitively-driven and rooted in objective information about specific product features (Drolet, Simonson, & Tversky, 2000;
In contrast, hedonic assessments are grounded in emotional assessments of subjective content, particularly as consumer involvement increases (Park & Moon, 2003).

A notable feature of consumer-generated content such as eWOM is its subjective basis. A series of experiments conducted by Sen and Lerman (2007) shed light on this phenomenon. In these studies, consumers were asked to read a series of textual, web-based product reviews for utilitarian and hedonic products. For hedonic products, negative reviews were attributed to the subjective bias of consumer reviewers. In contrast, unfavorable reviews of utilitarian products were attributed to product characteristics.

Integrating these insights, we see that consumers a) utilize subjective content to form their judgments of hedonic products and b) believe that consumer-generated content regarding these products is more subjective. Thus, it stands to reason that consumer-generated text will evoke more attention in the case of hedonic products as opposed to utilitarian products. In the word-of-mouth literature, the utilitarian/hedonic distinction is often talked about in terms of luxury versus non-luxury brands (e.g., Berry, 1994; Patrick & Haugtvedt, 2009). In particular, luxury brands are often identified in terms of their hedonic attributes, while non-luxury brands are framed in terms of their utilitarian functionality. We adopt the luxury/non-luxury distinction here and suggest the following hypothesis:

Hypothesis 1: Attention to image versus text-based elements of eWOM will be qualified by the brand. Specifically, consumers will pay more attention to eWOM for a) image-based elements of non-luxury brands and b) text-based elements of luxury brands.

Further Qualifying Attention to Structural Elements: Word-of-Mouth Type

As mentioned previously, consumer-generated eWOM may include evaluative judgments (Buttle, 1998; Carl, 2008), non-evaluative information sharing (Richins & Root-Shaffer, 1988), and communal “buzz” (Dye, 2000). In contrast to traditional modes of advertising, eWOM is believed to be particularly influential due to its naturalistic source (Keller, 2007), the speed with which it is transmitted (Jurvetson, 2000; Mohr, 2001), and the ease with which it is retrieved (Bakos, 1997; Hoffman & Novak, 1997). Further distinction is drawn between positive word-of-mouth (PWOM), or favorable reviews of products and services, and negative word-of-mouth (NWOM), or unfavorable reviews (see Luo, 2009). Meanwhile, neutral word-of-mouth has an informational tone, yet can still be persuasive by virtue of facilitating consumer awareness.

Sen and Lerman (2007) note that consumers possess an attitudinal bias towards negative eWOM (see also Ahulwia & Shiv, 1997). This is because the relative scarcity of negative informational cues (in comparison to positive and neutral) makes them more salient, engendering increased attention (see Zajone, 1968). These authors subsequently demonstrate that subjects engaged more with negative reviews of hedonic (luxury) as opposed to utilitarian (non-luxury) products. It appears that this was due to the fact that negative reviews of luxury products were more expectation-inconsistent (and thus, attributed to rater biases) than negative reviews of utilitarian products, which were feasible, and thus attributed to product characteristics. Notably, the reviews presented to subjects were predominantly text-related. However, Sen and Lerman (2007) neither measured attention directly (i.e., behaviorally) nor included product images, which are frequently utilized by consumers in a social media context. Based on these findings, we suggest the following three-way interaction will occur:

Hypothesis 2: Attention to image versus text-based elements of eWOM will be qualified by message type. Specifically, consumers will pay more attention to eWOM text elements for luxury brands containing negative word-of-mouth (NWOM).

METHOD

To test these hypotheses, a within-subjects 2 x 3 x 2 factorial design experiment was conducted (see Table 1). Consumer attention was measured using eye-movement tracking across a series of consumer-generated eWOM social media pages with structural elements (text vs: image), message valence (negative vs: positive vs: neutral) and brand type (luxury vs: non-luxury) serving as the independent variables. Consumer research on attention typically utilizes

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand Luxury</td>
<td>24140.19</td>
<td>1</td>
<td>24140.19</td>
<td>3.22</td>
<td>.084</td>
</tr>
<tr>
<td>WOM Type</td>
<td>2464485.47</td>
<td>2</td>
<td>1230592.74</td>
<td>56.09</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>Text/Image</td>
<td>105010.71</td>
<td>1</td>
<td>105010.71</td>
<td>3.08</td>
<td>.090</td>
</tr>
<tr>
<td><strong>Interactions (2-Way)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luxury x WOM Type</td>
<td>214653.58</td>
<td>2</td>
<td>107326.79</td>
<td>8.31</td>
<td>.001**</td>
</tr>
<tr>
<td>Luxury x Text/Image</td>
<td>212809.33</td>
<td>1</td>
<td>212809.33</td>
<td>27.54</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>WOM Type x Text/Image</td>
<td>39119.05</td>
<td>2</td>
<td>19559.53</td>
<td>2.06</td>
<td>.137</td>
</tr>
<tr>
<td><strong>Interaction (3-Way)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luxury x WOM type x Text/Image</td>
<td>106449.26</td>
<td>2</td>
<td>53224.63</td>
<td>7.81</td>
<td>.001**</td>
</tr>
</tbody>
</table>
a within-subject design (e.g., Lohse, 1997; Pieters & Wedel, 2004) as a means of ascertaining differences in attention to various categories of stimuli by the average consumer.

**Participants.** A total of twenty-eight undergraduates enrolled at a Midwestern university in the US took part in this study. This sample was chosen based on research suggesting that eWOM is most likely to impact the purchase decisions of 15-24 year-olds (Wallace, Walker, Lopez, & Jones, 2011).

**Stimuli.** We constructed a series of web-pages containing consumer-generated product reviews using the popular social media platform Pinterest. This platform functions as a virtual pin board, enabling users to search the internet for images, “pin” them to thematic “boards” or pages, and add any comments they may have about the images they post. Other Pinterest users have access to these boards, and are free to “re-pin” the images and comments that they like. Pinterest stands apart from other social media platforms by juxtaposing images and text. Furthermore, one of its explicit emphases is consumer-driven eWOM, as evidenced by a recent survey (Boticca.com, 2012) indicating that consumers referred to product websites from Pinterest spent an average of $180. Facebook users, in contrast, spend an average of $85.

Numerous product categories were considered for this experiment. We chose to use restaurants based on recent evidence (Harris Interactive, 2006; cf. Allsop, Bassett, & Hoskins, 2007) that eWOM pertaining to restaurants is sought out and provided online more than any other product category. Specific representatives of luxury and non-luxury restaurants were chosen on the basis of previous research recommendations for studying the luxury/non-luxury distinction (see Arora, 2012). In this instance, Hyde Park Prime Steakhouse was selected to represent the luxury (hedonic) category, while Applebees was selected as a non-luxury (utilitarian) counterpart offering a comparable type of cuisine.

Separate product pages were constructed and pre-tested using Pinterest for each brand and designed to be viewed on a single computer screen, eliminating the need for subjects to scroll or navigate to other locations. The text for the positive, negative, and neutral eWOM stimuli were adapted from actual social media content and combined with product images corresponding with standard consumer-generated eWOM on Pinterest. Images chosen by consumers to coincide with product reviews included restaurant logos, the interior/exterior design of the restaurant, staff representatives, and prepared meals. Eight image/text pairs were selected for each of the two brand pages, resulting in a total of 16 image/text pairs that were randomized in order to minimize the influence of order effects. A pilot study (n = 56) was done to validate the brand type and message valence stimuli. Paired comparison t-tests indicated significant mean differences between the luxury ratings of Applebees and Hyde Park Prime Steakhouse (m_{Applebees} = 2.73, m_{HydePark} = 5.69, t = 9.86). For message valence, subjects were asked to evaluate the 16 text/image pairs and rate the extent to which they represent a) positive, b) negative, and c) neutral (e.g., restaurant hours, location) word-of-mouth using a 7-point Likert scale (e.g., 1 = not at all positive, 7 = very positive). Within-subject ANOVA results indicated that participants viewed PWOM combinations as more positive in nature than negative or neutral (F_{2, 22} = 256.09, p < .001). Consistent results were found for the negative (F_{2, 22} = 182.75, p < .001) and neutral (F_{2, 14} = 47.14, p < .001) text/image pairs.

**Dependent Measure.** Consumer attention was measured as the dependent variable and operationally defined as the total number of times participants fixated on pre-identified Areas of Interest (AOIs). Fixation frequency was measured using the Multi-Analysis of Psychophysiological and Performance Signals (MAPPs) eye-tracking system produced by eyesDx. To pinpoint consumer attention, each restaurant page was coded in terms of 16 AOIs (8 images, 8 text elements). Once relevant AOIs are identified, the eye-tracking software monitors and records participant eye movement and fixation frequency within the AOIs and data is subsequently aggregated for analysis.

**Procedure.** Subjects first completed a brief online survey designed to collect background information on each participant. Next, a script was read to participants introducing the eye-tracking software followed by an eye-movement calibration exercise. Calibration is an important step in eye-tracking research in that it syncs participant eye movements with the monitoring equipment (see Hornof & Halverson, 2002). Subjects were then instructed that the purpose of the study was to record their evaluation of a series of consumer-generated brand-related Pinterest pages. Participants proceeded to view the stimulus pages for two minutes per restaurant, with a ten second interval between pages. Following this phase of the experiment, subjects completed another brief online survey and were thanked for their participation.

**RESULTS**

The sample consisted of 7 males (25%) and 21 females (75%) with an average age of 24.9. Samples of 20-30 participants are fairly typical for eye-tracking, which a) are less susceptible to human error/bias due to their physiological basis and b) produce a large number of within-subjects data points (i.e., several thousand). Consistent with the pre-test, subjects rated Hyde Park Prime Steakhouse as more luxurious than Applebee’s (m_{Applebees} = 1.43; m_{HydePark} = 4.25; t = 7.20, p < .01). To rule out any brand preference bias, we assessed brand attitudes using a 7-item scale developed by Shamsdani, Stanaland, and Tan (2001). Differences in attitudes towards the two brands were non-significant (t = 0.309, p = .760).

**Hypothesis Testing.** Hypothesis 1 predicts a significant two-way interaction of structural elements and brand luxury. In particular, subjects were expected to fixate more frequently on a) the image-based elements of non-luxury brands and b) the text-based elements of luxury brands. A statistically significant interaction was found (F_{1, 7} = 27.54, p < .001). More specifically, participants paid more attention to image-based (m = 236.64 fixations) as opposed to text-based (m = 150.95 fixations) elements of non-luxury restaurant reviews. In contrast, attention to text-based elements (m = 218.24 fixations) was greater than attention to image-based elements (m = 203.26 fixations) for reviews of the luxury restaurant. Thus, hypothesis 2 was supported.

Hypothesis 2 predicts a significant three-way interaction of structural elements, brand luxury, and word-of-mouth type. Attention to text was expected to be pronounced for luxury brands that were evaluated with negative word-of-mouth. The results indicated a statistically significant three-way interaction (F_{2, 54} = 7.81, p = .001), such that attention to the text-based elements of luxury restaurant NWOM (m = 350.32 fixations) was greater than attention to the image-based elements of this condition (m = 256.86 fixations). In contrast, mean levels of attention to image-based elements as opposed to text-based elements was higher in each of the remaining conditions, as Figure 1 illustrates. Thus, hypothesis 3 was supported.

**DISCUSSION**

Consumer attention to structural aspects of eWOM is especially important within a social media context, where image and text-based elements of products differ and compete to get noticed. With the aid of behavioral eye-tracking, we have demonstrated empirically that notable attention-based differences do exist as a function of struc-
Our findings also provide support to the notion that consumer-generated eWOM should not be ignored by marketers in favor of product promotion opportunities that offer more direct control (Daugherty & Hoffman, in press; Graham & Havlena, 2007; Jepsen, 2006). Indeed, while numerous suggestions for leveraging eWOM have been made to marketing firms (see Mason, 2008) we add to this list by suggesting more structure-related strategies. An image-based example of this strategy is the recent “Show us your pizza” campaign launched by Dominos pizza (Alfs, 2013). Using social media, this company encouraged customers to take and post pictures of the pizzas that were delivered to them. The results of our study predict that positive images of pizzas (a utilitarian, non-luxury product) would be more attended to than text-based reviews. The reported success of this campaign (Alfs, 2013) appears to support this notion.

Framed in terms of making a difference, our findings have a number of intriguing implications. As alluded to above, social media users that advocate difference-making could be encouraged by marketers to communicate their perspective in strategic ways that maximize the odds of their content being attended to. For instance, difference-making can be framed in hedonic and luxury-based terms (e.g., a formal benefit dinner) or in more utilitarian terms (e.g., acts of community service). Difference-makers could also be prompted to post compelling pictures (image-based stimuli), moving stories (text-based stimuli), or both as a function of the difference-making activity and how it is perceived (i.e., hedonic or utilitarian). Final-
ly, negative word-of-mouth could be leveraged as a valuable tool to bring attention to important issues where difference-making is needed. For instance, our findings suggest that writing a negative review of luxury items on the basis that their business practices are unethical would be attended to more than a negative image of the brand. In short, it may be possible to leverage structural attention dynamics to optimize the impact of difference-making solicitations through social media.

Despite the strength of our findings, some important limitations merit mentioning. By using Pinterest as the platform for our study, we held constant other structural elements of eWOM presentation such as size (Pieters & Wedel, 2004). Additional work is needed to assess the unique contribution of these elements to attention within a social media environment. Related to this, an important aspect of online environments is the ability to navigate between social media pages. We limited participant attention to four contrived Pinterest pages so that Areas of Interest could be coded, a key component of eye-tracking research. Future studies could utilize other approaches to coding AOIs that give participants control over what they view.

**CONCLUSION**

The growth of social media and eWOM make it more important than ever to understand dynamics of consumer attention. As such, existing models of eWOM are incomplete without taking into account the consumer attention process within social media. Although much remains to be done, our work provides support for a long-held supposition; that communication effectiveness is not just bound by the content of a message, but also delivery effectiveness. Presumably, consumers invest their time and energy into generating eWOM on social media because they desire to make a difference by influencing product awareness and decisions made by other consumers. The managerial implications of our results are immediate as the ability to build brand relationships via social media represents the future of marketing. We call for additional research designed to aid consumers and product marketers alike in utilizing structural aspects of eWOM to their fullest potential.

**REFERENCES**


Daugherty, Terry and Ernst Hoffman (in press), “eWOM and the Importance of Capturing Consumer Attention within Social Media,” *Journal of Marketing Communications*.


