7-1-2013

# On Top Of The World, Down In The Dumps: Text Mining The Emotionality Of Online Consumer Reviews

Maria Madlberger
*Webster University Vienna, Vienna, Austria*, madlberger@webster.ac.at

Makoto Nakayama
*DePaul University, Chicago, Illinois, United States*, mnakayama@cdm.depaul.edu

# ON TOP OF THE WORLD, DOWN IN THE DUMPS: TEXT MINING THE EMOTIONALITY OF ONLINE CONSUMER REVIEWS

Madlberger, Maria, Webster University Vienna, Berchtoldgasse 1, A-1220 Vienna, Austria, madlberger@webster.ac.at

Nakayama, Makoto, DePaul University, 243 South Wabash Avenue, Chicago, IL 60604, United States, MNakayama@cdm.depaul.edu

## Abstract

*Emotionality of online reviews can reveal much knowledge about product perceptions and popularity. However, current emotion-related analyses of online reviews are mainly focused on sentiment and valence analysis. Based on the theory of discrete emotions, we differentiate between six different emotions (anger, disgust, fear, joy, sadness, surprise) and posit that they appear differently depending on product types. To assess emotions, we introduce two key measurement dimensions: emotional breadth (word variety) and emotional depth (total word counts). We hypothesize that the density of emotional breadth and depth is higher for hedonic goods than for utilitarian goods as well as that the densities of the six emotions differ between these product types. To test the hypotheses, we collected 10,087 reviews of 30 printers and 30 music CDs at Amazon.com. The text mining analyses on these reviews show that reviews on music CDs are much more emotional than those on printers. Joy is the dominant emotion in both product categories. However, the emotions of disgust (emotional breadth) and sadness (emotional depth) are more prominent for printers than for music CDs. Implications and future research avenues are discussed along with the challenges of textual keyword analyses on discrete emotions.*

*Keywords: Text mining, online consumer reviews, theory of discrete emotions, emotional breadth, emotional depth, emotional density, utilitarian goods, hedonic goods*

# 1 Introduction

Understanding emotions in online reviews can largely improve the comprehension of mechanisms that make certain products more popular than others (Kim and Gupta 2012). It is useful for determining the "emotional load" of different products or product categories. In the context of online reviews, an analysis of emotional terms allows investigating whether one product type is indeed associated with more intensive emotions than another one. Furthermore, knowledge of collective emotions helps better understand how products are being rated and what factors increase or decrease products' popularity as well as their perceived attributes (Garcia and Schweitzer 2011).

To date, text mining of product evaluations in online consumer reviews is largely focused on sentiment and valence analysis that distinguishes between positive and negative remarks (Das and Chen 2007; Pang and Lee 2008). However, similar to word of mouth, online reviews are also characterized by emotionality (Kim and Gupta 2012). Individuals who write reviews look for textual or symbolic expressions that replace the missing possibilities to express emotions non-verbally (Provine et al. 2007; Xu et al. 2007). Some researchers argue that emotional product evaluations are less useful than ones with rational appeal as the former are subjective, while the latter are objective (Park and Lee 2008). This is, however, not clearly evidenced. In contrast, empirical research conducted by Li and Zhan (2011) shows that consumers appreciate strong positive emotional reviews. This finding contradicts the expectation that usually objective opinions are preferred.

Emotions provide more differentiated insights than positive or negative valence, as discrete emotions on the same valence influence judgments differently (Lerner and Keltner 2000). Research on emotions and affect in online reviews started just recently and is far from being comprehensive. Exceptions are the works by Kim and Gupta (2012), Park and Lee (2008), and Sen and Lerman (2007) who investigate the impact of emotional expressions in online reviews on product evaluations. However, all studies consider emotions on the basis of valence. Kim and Gupta (2012) consider positive and negative valence of reviews, whereas Park and Lee (2008) analyze a particular selection of emotional reviews, i.e., short, highly subjective texts. Similarly, Sen and Lerman (2007) investigate whether positive or negative reviews are considered more useful and reveal that significant differences between utilitarian and hedonic products occur. Further, Pollach (2006) analyzes the amount of emotive appeals, appeals to reason, and reports on experience in online reviews

In order to contribute to a finer-granulated analysis of single emotions rather than positive or negative sentiments in online reviews, we draw on discrete emotions theory (Izard 1977; Izard and Malatesta 1987), a theoretical approach that distinguishes several basic emotions which are independent from each other. In doing so, we provide a refined analysis of occurrence of emotional terms in product reviews that supplements existing research on sentiment analysis and positive/negative affect.

Methodologically, our approach is text mining-based and similar to that of Garcia and Schweitzer (2011), however it is done on a finer scale as being based on assigning emotional terms to six discrete emotions rather than valence. As a consequence, our research is faced with challenges that go beyond valence identification and are particularly driven by assigning equivocal terms (e.g., to get) to the corresponding "right" emotion. Thus our analysis provides a fine-granulated approach of lexicon-based text mining for emotional terms and develops a set of rules to cope with equivocal terms that are not sufficiently handled in available lexicons.

In order to demonstrate a useful application context, we apply our text mining approach to an analysis of emotion intensity of online reviews in the context of different product categories. In the initial phase of the development of our research instrument, we focus on two product types that are considered different in terms of their affective characteristics in literature, i.e., a utilitarian and a hedonic good. The paper is organized as follows: In the next section a literature review on emotions in consumer behavior as well as text mining of emotions is provided. Section three discusses the hypotheses development. In section four, the text mining method is described. This section also raises challenging issues of identifying emotional terms. Section 5 presents the results of the hypotheses tests and section 6 discusses the implications from the findings.

# 2 Theoretical Background

## 2.1 Research on emotions in consumer behavior

Since the 1970s, a large number of studies show that emotions play a key role in many facets of consumer behavior (see Laros and Steenkamp 2005 for an overview). Within the context of product evaluation, the role of emotions is twofold: First, consumers' judgments of products are influenced by their current mood (Gorn et al. 1993; Kim et al. 2010; Pham 1998), whereby particularly negative feelings can result in more negative product evaluations. Second, the product itself as well as a product's emotional claims can evoke emotions that are being expressed in product evaluations. Research revealed that the more product claims are congruent with individuals' feelings, the more positively products are being rated (Martin et al. 1997). Product evaluation itself can take place at different levels of emotionality (Kim and Gupta 2012; Pang and Lee 2008; Pollach 2006). In the context of product evaluation, Kim et al. (2010) investigate the role of congruency between a product's emotional appeal (e.g., adventurousness) and individuals' feelings about the product. Similarly, the congruence between salient goals of a product consideration and specific emotions impact product evaluation. Thus achievement-related emotions are more informative for a product evaluation if achievement goals are dominating; the same is true for congruence between protection-related emotions and goals (Bosmans and Baumgartner 2005). Furthermore, the representation of product information influences how emotions are related to product evaluation. Verbal product information leads to dependence of emotions' impact on evaluation from relevance of hedonic or utilitarian evaluation factors. However, if the product is seen before verbal information is given, judgments are made based on the visual impression rather than hedonic or utilitarian criteria (Yeung and Wyer 2004). Emotions are also important factors in adoption and evaluation of innovative products, as Wood and Moreau (2006) show. Their empirical research reveals that emotions that emerge during a learning process on innovations influence product evaluation.

## 2.2 Categorizations of emotions

Researchers repeatedly raised the issue of classifying emotions and discussed how broadly emotions should be defined (Laros and Steenkamp 2005). Thus there are several complementary approaches depending on the context of interest. Some scholars analyze emotions on an aggregate level as positive and negative affect and thus focus on the polarity of affects (e.g., Oliver 1993). To analyze mood, emotions are considered dimensions on a continuum from positive to negative affect (Tellegen and Watson 1999). The theory of discrete emotions (Izard and Malatesta 1987) is grounded in the research on facial expression of emotions by Ekman et al. (1969). It differentiates between several discrete emotions, thus identifying a larger number of emotions that are not interdependent (e.g, Nabi 2003). Since Izard (1977) discussed ten primary emotions, numerous studies investigated the role of discrete emotions in various contexts. Examples of studies in the context of consumer behavior are Lerner and Keltner (2000), Richins (1997), Nyer (1997), and Ruth et al. (2002). Dobele et al. (2007) investigated the rule of discrete emotions in an e-commerce context, that is, viral marketing. Many researchers distinguish basic or primary emotions that are considered building blocks of further, nonbasic emotions (Ortony and Turner 1990). Despite large differences in basic emotion categorizations (see Ortony and Turner 1990 for a review), six emotions are widely accepted as basic emotions. These emotions are anger, disgust, fear, joy, sadness, and surprise (Oatley and Jenkins 1996; Parrot 2001; Plutchik 1980). Discrete emotions are particularly relevant in research on the expression of emotions in a verbal way that seeks to elaborate lexical semantic approaches for emotional words (Ortony et al. 1987; Clore et al. 1987). Verbal expression of emotions is complementary to the non-verbal one and provides particularly insights into an individual's emotional state (Fussell 2003).

### 2.3 Text mining approaches to identify emotional terms

Text mining is a significant application area that yields useful analysis results in many contexts. Thus various studies use text mining to conduct sentiment analysis, opinion mining (Baccianella et al. 2009), or to identify emotional terms in texts (Strapparava and Mihalcea 2007; Strapparava and Mihalcea 2008; Strapparava et al. 2006). Text mining on emotions builds upon the verbal expression of emotions. With increasing importance of text mining, the analysis of emotions that makes finer differentiations than valence or arousal is needed (Valitutti et al. 2004; Strapparava and Valitutti 2004). To capture emotional expressions in texts, not only the emotional states themselves, which represent the primary emotions, need to be considered, but also terms that are associated with emotion, such as causes or emotional behavior (Ortony et al. 1987) that are referred to as direct and indirect affective words (Strapparava et al. 2006). Text mining-based research that investigates emotions in online reviews is currently focused on valence and sentiment analysis. Garcia and Schweitzer (2011) collect emotional terms in online reviews based on two lexicons (SentiStrength and ANEW), but measure the emotions only in terms of their positive/negative valence. The study by Xia and Bechwati (2008) investigates the role of a person's affective intensity as a personal trait and the emotionality of a review for the overall review impact and purchase intention. However, at present, no empirical study addresses an in-depth analysis of the occurrence of single emotions in online reviews independent from review length or structure as well as valence.

## 3 Hypotheses

The nature of different products varies considerably. A widely accepted distinction is made between utilitarian and hedonic goods (Hirschman and Holbrook 1982; Strahilevitz and Myers 1998). The roots of this distinction lie in different types of consumption. Utilitarian consumption relates to satisfying a need or accomplishing a task, thus pursuing practical goals and satisfying necessities. On the other hand, hedonic consumption is pleasure-oriented and driven by fun, fantasy, and the desire to experience sensual pleasure (Strahilevitz and Myers 1998). As each product comprises varying degrees of utilitarian and hedonic characteristics, classifications between utilitarian and hedonic goods can be made (Hirschman and Holbrook 1982). Although products can hardly be classified as purely utilitarian or hedonic (e.g., cars), it can be assumed that for each product one aspect is dominating (Dhar and Wertenbroch 2000; Okada 2005). Product evaluations are largely driven by the product nature (Batra and Ahtola 1990). For utilitarian products, evaluations are mainly cognitive-driven, goal-oriented, and instrumental (Strahilevitz and Myers 1998) whereas for hedonic products, evaluation criteria are closely related to the consumption experience (Batra and Ahtola 1990). The nature of products does not only influence the evaluation criteria of products, but also the usefulness of sentiments in product reviews. For utilitarian products, negative reviews are considered more useful as the negative expressions are rather attributed to the product than the reviewer. For hedonic products, the opposite is true, i.e., negative reviews are more attributed to the reviewer and thus considered less useful (Sen and Lerman 2007).

Thus we conclude that there are differences in quantity (number of emotional terms) and quality (number of terms related to the basic emotions) of product reviews for utilitarian and hedonic products. To measure these dimensions of emotionality, we consider breadth, depth, and density of emotions in product reviews. The concepts of breadth and depth are based on the work by De Luca and Atuahene-Gima (2007) who introduced dimensions of breadth and depth in the context of knowledge. In analogy to this conceptualization of knowledge, the breadth of emotions measures the variety of emotional words used in reviews. The depth of emotions is defined as the total number of emotional words used in reviews. Finally, we define the density of emotions as the proportion of emotional words over all words in reviews to standardize the results for variations in review length. The notion of density is important because emotional words are embedded in all the other words. The impact of one emotional word should intuitively be assumed not the same between a 10-word review and a 100-word review (provided that there are any other emotional words). Thus the breadth density of emotions refers to the variety of emotional words relative to the total number of words while the depth density of emotions is the total number of emotional words relative to the total number of words in the reviews.

Text mining-based sentiment analysis acknowledges the fact that sentiments need to be considered in the context of different product features. Thus researchers who conduct sentiment analysis in online reviews perform their analyses usually in several steps, starting with extraction of product features, identifying positive or negative sentiments on these features, and analyzing the results (Hu and Liu 2004a; Hu and Liu 2004b; Jindal and Liu 2006; Popescu and Etzioni 2005; Wei et al. 2010). This approach stresses the need to take individual differences between products into account. In our study, we acknowledge differences in the nature of products by assuming that there are differences in evaluation of hedonic and utilitarian products. Given the larger relevance of affective dimensions in hedonic consumption (Hirschman and Holbrook 1982), we assume that emotional issues will be raised more in reviews of hedonic products. Bickart and Schindler (2002) show that hedonic-task reviews contain more emotional words than reviews examined for utilitarian tasks. Therefore we hypothesize:

H1: The density of emotional terms in product reviews is higher for hedonic goods than for utilitarian goods.

Differentiating between the breadth and depth densities, we specify:

H1a: The breadth density of emotions in product reviews is higher for hedonic goods than for utilitarian goods.

H1b: The depth density of emotions in product reviews is higher for hedonic goods than for utilitarian goods.

Consumer research revealed that different types of goals are associated with utilitarian and hedonic goods. In case of utilitarian goods, consumers rather pursue prevention goals, i.e., goals of maximizing safety and responsibility. For hedonic goods, they rather seek to achieve promotion goals which refer to accomplishments and aspirations (Higgins 1997). Product evaluations are influenced by these goals consumers pursue when purchasing a product (Chernev 2004). As a consequence, utilitarian and hedonic goods evoke different emotions. As an empirical study shows, negative experiences mainly cause feelings of anger in case of utilitarian goods while they rather cause dissatisfaction with hedonic goods (Chitturi et al. 2008). Therefore we contend:

H2: The densities of the six basic emotions differ between reviews of utilitarian goods and hedonic goods.

Again differentiating between the breadth and depth densities, we hypothesize:

H2a: The densities of emotional breadth of the six basic emotions are different between reviews of utilitarian goods and hedonic goods.

H2b: The densities of emotional depth of the six basic emotions are different between reviews of utilitarian goods and hedonic goods.

# 4 Method

This study uses textual keywords to detect the six basic emotions (anger, disgust, fear, joy, sadness, and surprise) in reviews (Oatley and Jenkins 1996; Parrot 2001; Plutchik 1980). We first scanned online reviews and counted the breadth and depth of these keywords in consumer reviews for utilitarian and hedonic goods. Then we conducted statistical analyses to investigate the differences between the emotional levels of reviews for the two goods. The Shapiro-Wilk tests indicate that 6 (25%) out of the 24 emotional density variables are statistically non-normal. We then used Mann-Whitney tests to supplement t-test for validating the emotional level differences.

### 4.1 Data set

We used Amazon.com consumer reviews for the study. We selected printers as a utilitarian good and music CDs as a hedonic good (Sen and Lerman 2007). The number of reviews varies significantly from product to product even within the same product category. Thus we focused on reviews for the bestselling printers and music CDs. To assure statistical stability, we screened out the products whose numbers of reviews were less than 30. To test the hypotheses, we collected the first 30 products that met the criteria from the bestselling list. These reviews were collected in October 2012. There were 5,028 reviews for 30 printers, and 5,059 reviews for music CDs (excluding video-only reviews). The average words per review are 179.0 (printers) and 136.1 (music CDs).

### 4.2 Emotional words

The definition of emotional words (and phrases such as “be on cloud nine”) was defined by WordNet-Affect (Strapparava and Mihalcea 2008; Strapparava et al. 2006), which is an extension of the online lexical database WordNet (Miller 1995). The current version of WordNet lists 155,287 nouns, verbs, adjectives and adverbs[1]. The particular version of WordNet-Affect we used was the SemEval-2007 “Affective Text” (Strapparava and Mihalcea 2007)[2]. This dataset defines 1,536 words for the six emotional categories: anger (318 words), disgust (72 words), fear (208 words), joy (539 words), sadness (309 words), and surprise (90 words) (Bobicev et al. 2010). However, WordNet-Affect is considered “a small lexical resource” compared to WordNet; it facilitates the analysis of emotions (Torii et al. 2011). We regard WordNet-Affect as the baseline emotional keyword sets rather than a comprehensive one.

### 4.3 Text mining analyses

We counted each emotional word or phrase appearing in reviews. The WordNet-Affect words include verbs and adjectives. Given WordNet-Affect lists only their basic form, we created a list of verb and adjective variations based on the WordNet-Affect dataset. The list includes: (a) third-person singular, past tense, past perfect tense and gerund forms of verbs, (b) verb exceptions (e.g., get-got-gotten), (c) comparative and superlative forms of adjectives (e.g., low-lower-lowest), and (d) adjective exceptions (e.g., good/well-better-best, happy-happier-happiest). Each noun was counted as the same word whether it appeared in singular or plural form. We referenced the WordNet definition of verb, adjective, and noun exceptions.

### 4.4 Word-sense considerations

One important limitation of the simple keyword approach is the lack of semantic recognitions of emotional words (Wu et al. 2006). This directly relates to our ability to determine the sense of a word in a particular linguistic context, or word-sense induction (WSI) (Bordag 2006). For example, an “angry” word, “get” (e.g., “His lying really gets me.”), is most likely used as a non-angry word (e.g., “I decided to get this CD.”). Ignoring word senses, we are then likely to overestimate the emotional levels. While experimental WSI learning systems are reported (Manandhar et al. 2010), commercial WSI software is not readily available. However, our preliminary analyses of the reviews show that 10% of emotional words account for 90% of usage share. Indeed, half of the emotional categories see their top 20 words having a 95% or higher usage share. Therefore, we observed that a few problematic words within the top 20 lists exist. Thus, to minimize the overestimation of emotions, we adjusted the word counts according to their word sense usage frequencies.

These problematic words were: fit, get at, get to, scene, score, and spite (anger); like, look for, move, close, near, catch, kid, concern, and worry (joy); blue, down, low, and dark (sadness); get and

[1] http://wordnet.princeton.edu/wordnet/man/wnstats.7WN.html
[2] http://www.cse.unt.edu/~rada/affectivetext/

wonder (surprise). We first scanned the entire reviews to assess how these words are used. This reveals that some emotional words are used repeatedly in certain linguistic structural contexts for their emotional sense(s). For example, the emotional word "like" (find enjoyable, be fond of) is used as a verb in the context of <person(s)> + [really] + like, such as I/you/we (really) like/liked the product. We calculated the frequency of such structures appearing in each good category. If structural cues are not applicable, we collected the relative sense frequency data on such words from WordNet (Mihalcea et al. 2004) and we used the frequency data to estimate relative sense usage ratios (frequencies of applicable senses divided by frequencies of all senses). Also we considered the particular context of our product reviews. For instance, the word blue is used exclusively as a color for printers. Music CD reviews had two instances of "blue" as depressed or low-spirited out of 57 instances (the sense frequency of 2/57 = 3.5%); the rest are part of song titles (e.g., Blue October) and non-emotional context (e.g., out of blue). Taken these considerations, we dropped the words whose relative emotional sense frequency is less than 1%. These dropped words were: fit, get at, get to, scene, score, and spite (anger); move*, close*, near*, catch, and worry (joy); blue*, low* and dark* (sadness); get (surprise) (*: dropped only for printers). The counts of other problematic words were adjusted by the relative emotional sense frequency estimates. To minimize the degree of subjectivity, both authors conducted the word sense evaluation independently from each other, with a high degree of inter-rater reliability.

# 5 Results

The results are summarized in Figure 1 and clearly indicate particular patterns of emotionality across the product categories. The y-axes show the average breadth/depth densities of each emotion in each product category. Joy is by far dominating in both product categories as well as breadth and depth densities. The second and third most important emotions are sadness and surprise. There is a clear difference between emotional breadth and depth densities when it comes to the emotions of anger, fear, and disgust, as they are considerably higher in terms of emotional breadth density.

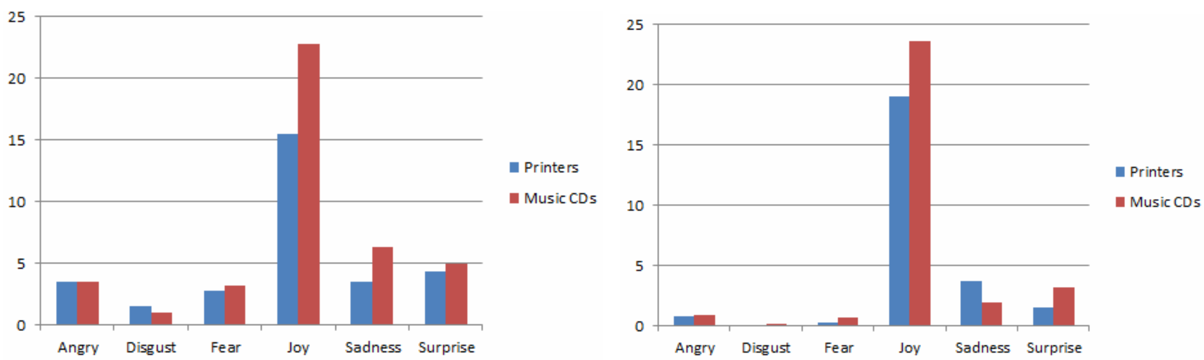


*Figure 1. Graphical representations of emotional breadth (right) and depth (left) densities*

In Table 1, both breadth and depth of total emotional density for music CDs were higher than those for printers. In breadth, music CD reviews are 35.9% more emotional than printers (t = -11.03, df = 49, p = .000). In depth, music CD reviews are 19.8% more emotional than printers (t = -4.11, df = 49, p = .000). These results support H1a and H1b.

| Emotional Intensity | Breadth Mean | S.D. | Depth Mean | S.D. |
|---|---|---|---|---|
| Printers | 31.10 | 2.83 | 25.47 | 3.53 |
| Music CDs | 41.87 | 4.42 | 30.52 | 5.61 |
| The unit is per 1,000 words. | | | | |

*Table 1. Total emotional densities for the two good categories*

The emotional breadth densities of disgust, joy, sadness, and surprise are statistically different between printers and music CDs (Table 2). This supports H2a. For the emotional depth density, disgust, fear, joy, sadness, and surprise are different between the two good types (Table 3). Thus, H2b is supported.

| Emotional Intensity (Breadth) | Anger | Disgust | Fear | Joy | Sadness | Surprise |
|---|---|---|---|---|---|---|
| Printers | 3.49 | 1.50 | 2.79 | 15.48 | 3.51 | 4.33 |
| Music CDs | 3.49 | 1.00 | 3.23 | 22.81 | 6.35 | 4.98 |
| Statistical Difference (t-test) | $t = -0.00$ $df = 52$ $p = 0.998$ | $t = 2.26$ $df = 56$ $p = 0.028$ | $t = -1.11$ $df = 56$ $p = 0.270$ | $t = -11.62$ $df = 41$ $p = 0.000$ | $t = -7.45$ $df = 40$ $p = 0.000$ | $t = -2.09$ $df = 54$ $p = 0.042$ |
| Statistical Difference (Mann-Whitney U test) | 427.00 $p = 0.734$ | 269.00 $p = 0.007$ | 358.00 $p = 0.337$ | 5.00 $p = 0.000$ | 79.00 $p = 0.000$ | 335.00 $p = 0.089$ |

*Table 2. Total emotional breadth densities for the two good categories*

| Emotional Intensity (Depth) | Anger | Disgust | Fear | Joy | Sadness | Surprise |
|---|---|---|---|---|---|---|
| Printers | 0.79 | 0.05 | 0.32 | 19.04 | 3.74 | 1.52 |
| Music CDs | 0.88 | 0.18 | 0.64 | 23.61 | 1.99 | 3.22 |
| Statistical Difference (p-value) | $t = -0.58$ $df = 39$ $p = 0.568$ | $t = -4.13$ $df = 37$ $p = 0.002$ | $t = -3.55$ $df = 37$ $p = 0.001$ | $t = -3.95$ $df = 47$ $p = 0.000$ | $t = 6.60$ $df = 53$ $p = 0.000$ | $t = -7.23$ $df = 39$ $p = 0.000$ |
| Statistical Difference (Mann-Whitney U test) | 403.50 $p = 0.492$ | 225.50 $p = 0.001$ | 210.50 $p = 0.000$ | 239.00 $p = 0.002$ | 99.50 $p = 0.000$ | 72.00 $p = 0.000$ |

*Table 3. Total emotional depth densities for the two good categories*

Interestingly, the disgust emotional breadth was higher for printers than music CDs whereas its depth counterpart was higher for music CDs than printers. Furthermore, while music CD reviews have higher emotional depth levels on average, the emotional depth of sadness is actually higher for printers than for music CDs. Possible reasons for these phenomena are discussed in the next section.

# 6 Implications

There are several notable implications of the results. First, joy is the most dominant emotion for both utilitarian and hedonic goods. Joy stands out 3 to 5 times or even more than other emotions, in terms of emotional breadth and depth density. For usage shares, the lowest was 50% for emotional breadth on printers whereas the highest was 77% for emotional depth for music CDs. Thus, the bulk of emotions we look for in utilitarian goods relate to the pleasant emotion, joy. If we focus on the dominance of emotional levels, we can say the positive emotions overpower the negative ones. If marketers have only the sentiment analysis tools, the significance of such dominance can be tested for different goods. Second, besides joy, we notice that sadness and surprise follow joy in both breadth and depth of emotional densities. However, sadness depth is greater for printers than music CDs. Also printers' disgust was 50% higher than that of music CDs. One possibility is that consumers want to vent their sadness and disappointment more for utilitarian goods than for hedonic goods because they expect the functions of utilitarian goods to work as *required*. As hypothesized, consumers evaluate utilitarian and hedonic goods based on different goals and criteria (Chernev 2004; Chitturi et al. 2008; Higgins 1997). Our findings are largely consistent with these studies. If consumers' expectations are not met, they may feel something they paid for is unjustly delivered. In contrast, the values of hedonic goods are much more subjective, thus consumers would not feel disappointment the same way as they would for utilitarian goods. A lack of expectation fulfilment is rather perceived as "a loss of pleasure" of a hedonic good (Chitturi et al. 2008, p. 51). In contrast to the study by Chitturi et al. (2008), our findings indicate that the levels of angry emotion are similar between utilitarian and hedonic goods. The emotional words for fear are such words as horrible, terrible, afraid and awful. We consider these words are words of disappointment. Thus, third, looking across the intensities of the six emotions, we can expect subtle to significant differences in the six emotional levels for a variety of goods beyond the two goods we focused on.

Emotional appeals are used sporadically in product reviews (Pollach 2006). In our data set, just 3.1% and 5.0% of unique words (breadth) are emotional words in the reviews of utilitarian and hedonic goods, respectively. In terms of word counts (depth), 2.6% and 3.1% are emotional words for each good. In contrast, in our samples, about 16% of unique words are so-called stop words, most commonly used, semantically neutral words such as a, the, and you (Fox 1989). This gives us a perspective on emotional words. Therefore, fourth, emotional words are similar to vitamins in our diet in the sense that differences in small amounts can be differences between too little and too much. In the keyword approach, small proportions of keywords take about 90% of usage share. In our study, the top 20 emotional words in each category played major "emotional" roles. WSI corrections are essential in the keyword approach. For instance, we would have estimated at least 20% more on the emotional depth of joy for printer reviews. The emotion of surprise can be overestimated even more significantly, given the word "get" accounts for 74% and 48% of printer and music CD reviews. The sense for surprise is rare for the word "get" and the word is the fifth most frequently used verb following "be", "have", "do", and "say".[3] However, breadth is much less sensitive to WSI than depth because WSI affects directly word counts. Breath is not affected unless the emotional sense frequency of the word is negligible in the entire collection of reviews.

There are some limitations of this study. One significant issue is the validation of metrics. Emotional words appear sporadically in consumer reviews, thus a survey-based validation by asking individuals about how much emotion they feel with reviews is limited in contribution. More viable approaches could relate to objective emotion measurement, such as functional magnetic resonance imaging (fMRI) as suggested by Yoon et al. (2012). Our analysis of review emotionality is based on the emotional word counts. How these counts correspond on other definitions of emotionality would be a future research agenda. We used the SemEval-2007 "Affective Text" which does not include much colloquial vocabulary. For example, the word "crap" is used 18 times for printer reviews. The word might be regarded as sadness or anger. We look forward to using the enhanced version of "Affect Text". In relation to colloquial terms, the reviews are mainly written by American consumers. The differences between American and British English should be investigated further. We need to conduct follow-up studies as soon as reliable word-sense detection software is available.

# 7 Conclusion

Emotions play an important role in online consumer behavior, as emotions can accelerate purchasing decisions. While emotions were studied for advertisement and consumer behaviors in the past, the examination of emotions for online consumer reviews is relatively new and deserves more empirical studies. This research focused on reviews of hedonic and utilitarian goods. Their emotionality depends on the hedonic proportion of product attributes as we hypothesized. The sum total of different emotions is higher for the hedonic good than for the utilitarian good. However, the utilitarian good had higher emotionality in their reviews on emotions such as disgust and sadness. Thus, the study revealed that emotionality is not necessarily monolithic. However, the emotion of joy was the most dominant among the six emotions we investigated, thus validating the utility of sentiment analysis. From the theoretical standpoint, we showed that breadth and depth of emotionality can shed light on different aspects of emotions. Breadth is indicative of emotional variations within a certain emotional type whereas depth tells us how deep an emotion runs regarding the product purchase experience. While the emotional breadth of printer reviews is lower on sadness than that of music CDs, the emotional depth of sadness is higher for printers than for music CDs. Future studies should further investigate the roles of different emotions in online reviews for a variety of goods. We should also look into products

[3] http://oxforddictionaries.com/words/the-oec-facts-about-the-language

that are in the “middle" of the utilitarian-hedonic spectrum as this could deepen our understanding of the emotionality of consumer reviews on utilitarian and hedonic products.

## References

Baccianella, S., Esuli, A., and Sebastiani, F. (2009). Multi-facet rating of product reviews. In Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval, 461-472.

Batra, R. and Ahtola, O.T. (1990). Measuring the hedonic and utilitarian sources of consumer attitudes. Marketing Letters, 2 (2), 159-170.

Bickart, B. and Schindler, R.M. (2002). Expanding the scope of word of mouth: Consumer-to-consumer information on the Internet. Advances in Consumer Research, 29 (1), 428-430.

Bobicev, V., Maxim, V., Prodan, T., Burciu, N., and Angheluş, V. (2010). Emotions in words: Developing a multilingual wordnet-affect. Computational Linguistics and Intelligent Text Processing, 375-384.

Bordag, S. (2006). Word sense induction: Triplet-based clustering and automatic evaluation. In Proceedings of the 11th EACL, 137-144.

Bosmans, A. and Baumgartner, H. (2005). Goal-relevant emotional information: When extraneous affect leads to persuasion and when it does not. Journal of Consumer Research, 32 (3), 424-434.

Chernev, A. (2004). Goal-attributive compatibility in consumer choice. Journal of Consumer Psychology, 14 (1&2), 141-150.

Chitturi, R., Raghunathan, R., and Mahajan, V. (2008). Delight by design: The role of hedonic versus utilitarian benefits. Journal of Marketing, 72 (3), 48-63.

Clore, G.L., Ortony, A., and Foss, M.A. (1987). The psychological foundations of the affective lexicon. Journal of Personality and Social Psychology, 53 (4), 751-766.

Das, S.R. and Chen, M. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the Web. Management Science, 53 (9), 1375-1388.

De Luca, L.M. and Atuahene-Gima, K. (2007). Market knowledge dimensions and cross-functional collaboration: Examining the different routes to product innovation performance. Journal of Marketing, 71 (1), 95-112.

Dhar, R. and Wertenbroch, K. (2000). Consumer choice between hedonic and utilitarian goods. Journal of Marketing Research, 37, 60-71.

Dobele, A., Lindgreen, A., Beverland, M., Vanhamme, J., and van Wijk, R. (2007). Why pass on viral messages? Because they connect emotionally. Business Horizons, 50 (4), 291-304.

Ekman, P., Sorenson, E.R., and Friesen, W.V. (1969). Pan-cultural elements in facial displays of emotion. Science, 164 (3875), 86-88.

Fox, C. (1989). A stop list for general text. In Proceedings of the ACM SIGIR Forum, 19-21.

Fussell, S.R. (2003). The verbal communication of emotion: interdisciplinary perspectives. Lawrence Erlbaum Associates, Mahwah, NJ.

Garcia, D. and Schweitzer, F. (2011). Emotions in Product Reviews – Empirics and Models. In Proceedings of the Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE third International Conference on Social Computing (socialcom), IEEE, 483-488.

Gorn, G., Goldberg, M.E., and Basu, K. (1993). Mood, awareness, and product evaluation. Journal of Consumer Psychology, 2 (3), 237-256.

Higgins, T. (1997). Beyond pleasure and pain. American Psychologist, 52 (12), 1280-1300.

Hirschman, E.C. and Holbrook, M.B. (1982). Hedonic consumption: emerging concepts, methods, and propositions. Journal of Marketing 46, 92-101.

Hu, M. and Liu, B. (2004a). Mining and summarizing customer reviews. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, 168-177.

Hu, M. and Liu, B. (2004b). Mining opinion features in customer reviews. In Proceedings of the American Association for Artificial Intelligence (AAAI) Conference, 755-760.

Izard, C. (1977). Human Emotions. Plenum, New York.
Izard, C. and Malatesta, C. (1987). Perspectives on emotional development I: Differential emotions theory of early emotional development. In: Handbook of Infant Development, Osofsky, J.D. (ed.), Wiley, New York, 494-554.
Jindal, N. and Liu, B. (2006). Identifying comparative sentences in text documents. In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New Orleans, 244-251.
Kim, H., Park, K., and Schwarz, N. (2010). Will this trip really be exciting? The role of incidental emotions in product evaluation. Journal of Consumer Research, 36 (6), 983-991.
Kim, J. and Gupta, P. (2012). Emotional expressions in online user reviews: How they influence consumers' product evaluations. Journal of Business Research, 65 (7), 985-992.
Laros, F.J.M. and Steenkamp, J.B.E.M. (2005). Emotions in consumer behavior: A hierarchical approach. Journal of Business Research, 58 (10), 1437-1445.
Lerner, J.S. and Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. Cognition and Emotion, 14 (4), 473-493.
Li, J. and Zhan, L. (2011). Online persuasion: How the written word drives WOM. Evidence from consumer-generated product reviews. Journal of Advertising Research, 51 (1), 239-257.
Manandhar, S., Klapaftis, I.P., Dligach, D., and Pradhan, S.S. (2010). SemEval-2010 task 14: Word sense induction & disambiguation. In Proceedings of the 5th International Workshop on Semantic Evaluation, 63-68.
Martin, L.L., Abend, T., Sedikides, C., and Green, J.D. (1997). How would I feel if ...? Mood as input to a role fulfillment evaluation process. Journal of Personality and Social Psychology, 73 (2), 242-253.
Mihalcea, R., Tarau, P., and Figa, E. (2004). PageRank on semantic networks, with application to word sense disambiguation. In Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004), Geneva, Switzerland.
Miller, G.A. (1995). WordNet: a lexical database for English. Communications of the ACM, 38 (11), 39-41.
Nabi, R.L. (2003). Exploring the framing effects of emotion: Do discrete emotions differentially influence information accessibility, information seeking, and policy preference? Communication Research, 30 (2), 224-247.
Nyer, P.U. (1997). A study of the relationships between cognitive appraisals and consumption emotions. Journal of the Academy of Marketing Science, 25 (4), 296-304.
Oatley, K. and Jenkins, J.M. (1996). Understanding Emotions. Blackwell Publishers, Boston.
Okada, E.M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. Journal of Marketing Research, 42 (1), 43-53.
Oliver, R.L. (1993). Cognitive, affective, and attribute bases of the satisfaction response. Journal of Consumer Research, 20 (3), 418-430.
Ortony, A., Clore, G.K., and Foss, M.A. (1987). The referential structure of the affective lexicon. Cognitive Science, 11 (3), 341-364.
Ortony, A. and Turner, T.J. (1990). What's basic about basic emotions? Psychological Review, 97 (3), 315-331.
Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2 (1-2), 1-135.
Park, D.-H. and Lee, J. (2008). eWOM overload and its effect on consumer behavioral intention depending on consumer involvement. Electronic Commerce Research and Applications, 7 (4), 386-398.
Parrot, W. (2001). Emotions in Social Psychology. Psychology Press, Philadelphia.
Pham, M.T. (1998). Representativeness, relevance, and the use of feelings in decision making. Journal of Consumer Research, 25 (2), 144-159.
Plutchik, R. (1980). Emotion: A Psychoevolutionary Synthesis. Harper & Row, New York.

Pollach, I. (2006). Electronic word of mouth: A genre analysis of product reviews on consumer opinion web sites. In Proceedings of the 39th Hawaii International Conference on Systems Sciences (HICSS-39), IEEE, Waikoloa, HI, 51c-51c.
Popescu, A.-M. and Etzioni, O. (2005). Extracting product features and opinions from reviews. In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, 339-346.
Provine, R.R., Spencer, R.J., and Mandell, D.L. (2007). Emotional expression online: Emoticon punctuate website text messages. Journal of Language and Social Psychology, 26 (3), 299-307.
Richins, M.L. (1997). Measuring emotions in the consumption experience. Journal of Consumer Research, 24 (2), 127-146.
Ruth, J.A., Brunel, F.F., and Otnes, C.C. (2002). Linking thoughts to feelings: investigating cognitive appraisals and consumption emotions in a mixed-emotions context. Journal of the Academy of Marketing Science, 30 (1), 44-58.
Sen, S. and Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the Web. Journal of Interactive Marketing, 21 (4), 76-94.
Strahilevitz, M. and Myers, J.G. (1998). Donations to charity as purchase incentives: How well they work may depend on what you are trying to sell. Journal of Consumer Research, 24 (4), 434-446.
Strapparava, C. and Mihalcea, R. (2007). Semeval-2007 task 14: Affective text. In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Prague, Czech Republic.
Strapparava, C. and Mihalcea, R. (2008). Learning to identify emotions in text. In Proceedings of the The 2008 ACM Symposium on Applied Computing, ACM, Ceará, Brazil, 1556-1560.
Strapparava, C. and Valitutti, A. (Year). WordNet-Affect: an affective extension of WordNet. In Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004), Lisbon, 1083-1086.
Strapparava, C., Valitutti, A., and Stock, O. (2006). The affective weight of lexicon. In Proceedings of the International Conference on Language Resources and Evaluation, Genoa, Italy, 423-426.
Tellegen, A. and Watson, D. (1999). On the dimensional and hierarchical structure of affect. Psychological Science, 10 (4), 297-306.
Torii, Y., Das, D., Bandyopadhyay, S., and Okumura, M. (2011). Developing Japanese WordNet affect for analyzing emotions. In Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, Association for Computational Linguistics, Portland, Oregon, 80-86.
Valitutti, A., Strapparava, C., and Stock, O. (2004). Developing affective lexical resources. Psychology Journal, 2 (1), 61-83.
Wei, C.-P., Chen, Y.-M., Yang, C.-S., and Yang, C. (2010). Understanding what concerns consumers: A semantic approach to product feature extraction from consumer reviews. Information Systems & e-Business Management, 8 (2), 149-167.
Wood, S.L. and Moreau, C.P. (2006). From fear to loathing? How emotion influences the evaluation and early use of innovations. Journal of Marketing, 70 (3), 44-57.
Wu, C.H., Chuang, Z.J., and Lin, Y.C. (2006). Emotion recognition from text using semantic labels and separable mixture models. ACM Transactions on Asian Language Information Processing (TALIP), 5 (2), 168-183.
Xia, L. and Bechwati, N.N. (2008). Word of mouse: The role of cognitive personalization in online consumer reviews. Journal of Interactive Advertising, 9 (1), 3-13.
Xu, L., Yi, C., and Xu, Y. (2007). Emotional expression online: The impact of task, relationship and personality perception on emoticon usage in instant messenger. In Proceedings of the 11th Pacific-Asia Conference on Information Systems.
Yeung, C.W.M. and Wyer, R.S.J. (2004). Affect, appraisal, and consumer judgment. Journal of Consumer Research, 31 (2), 412-424.
Yoon, C., Gonzalez, R., Bechara, A., Berns, G.S., Dagher, A.A., Dubé, L., Huettel, S.A., Kable, J.W., Liberzon, I., and Plassmann, H. (2012). Decision Neuroscience and Consumer Decision Making. Marketing Letters, 23 (2), 473-485.