

# Evaluating Emotions in Mobile Application Descriptions: Sentiment Analysis Approach

*Completed Research Paper*

**Supavich (Fone) Pengnate**  
University of Wisconsin – Stevens Point  
fone.pengnate@uwsp.edu

**Dursun Delen**  
Oklahoma State University  
dursun.delen@okstate.edu

## Abstract

In this paper, we show how sentiment analysis can be used to detect emotions from mobile application descriptions which are persuasive text generated by software vendors. The analysis was conducted on a dataset containing 6,000 unique mobile apps. The results reveal that mobile app in different app category should use specific emotion words in the descriptions. This finding has interesting implications for practice especially for improving the chance that the app is downloaded or purchased.

## Introduction

In the online and information age, the Internet has enabled users to communicate and generate online content which usually in an exponential explosive rate. Examples of the content include messages users posted on social media websites, reviews and opinions from users, and product information posted by vendors. However, the majority of such content is typically in unstructured text format. This is difficult to decipher and analyze automatically, especially when evaluated tremendous amounts of data gathering from the Internet. Consequently, it is challenging for businesses to yield accurate information that is useful to them.

In order to analyze online content, especially in text format, there are several efficient methodologies a business can use to explore text, such as business intelligence, data mining, and natural language processing. However, in this paper, we adopted a recent technique called sentiment analysis to evaluate a large text dataset. In general, sentiment analysis is used to determine the attitude of a speaker or a writer with respect to some specific topics in text. The attitude can be in various forms such as opinions, evaluations, reviews, or emotional states of the author.

In this paper, mobile application (Android Apps) descriptions are analyzed using sentiment analysis and text mining approaches. The analysis was conducted to indicate salient emotions in app descriptions.

Although sentiment analysis has been recently explored in the text-mining literature, the focus of the existing studies has been primarily on the analysis of text generated by users, such as review, blog, an opinion. However, much less attention has been devoted to the analysis of text generated from the vendor's side, such as product description, advertising content, and webpage content. In the online commerce settings, vendor-generated content has been found to be a significant factor that influences customers' intentions (e.g., intention to use and to purchase (Chang et al. 1994)) since it provides the persuasive communication regarding the quality of products or services to the customers. Therefore, the primarily objective of this study is to investigate textual content generated from online vendors. Specifically, we investigate mobile app descriptions that are provided by a vendor through an online app store. Mobile app descriptions share the same mechanism as production descriptions which are written by software vendors with intent to persuade potential customers to download the apps.

Departing from the traditional text in the field of information processing, mobile app descriptions have a combination of linguistic characteristics. The language used in the descriptions is often quite specific to a particular domain (app categories, e.g., Games, Utilities, Finance, and Media & Video) and are generally written in persuasive communication fashion. Our approach to evaluate mobile app descriptions is to consider a description as a unit of text; however, we also consider the context of the app (app category),

each with its own topics and sentiments. This approach provides structural and useful information on the textual content at a fine-grain level. Our dataset contains 6,000 unique mobile apps in a total of 30 different app categories.

Writing product descriptions is an art. While most descriptions are arbitrarily written in a free-text format, it is difficult to assess and evaluate their effectiveness and efficiency in creating business value. This study proposes the application of sentiment analysis and text mining approaches to evaluate emotions conveyed in mobile app descriptions. This can help business develop or enhance strategies to improve user experience, which then leads to creating desired response. Specifically, the study aims to examine the following research questions:

1. How emotions in mobile app descriptions impact the number of downloads?
2. What are salient emotions that influence the number of downloads for a particular app category?

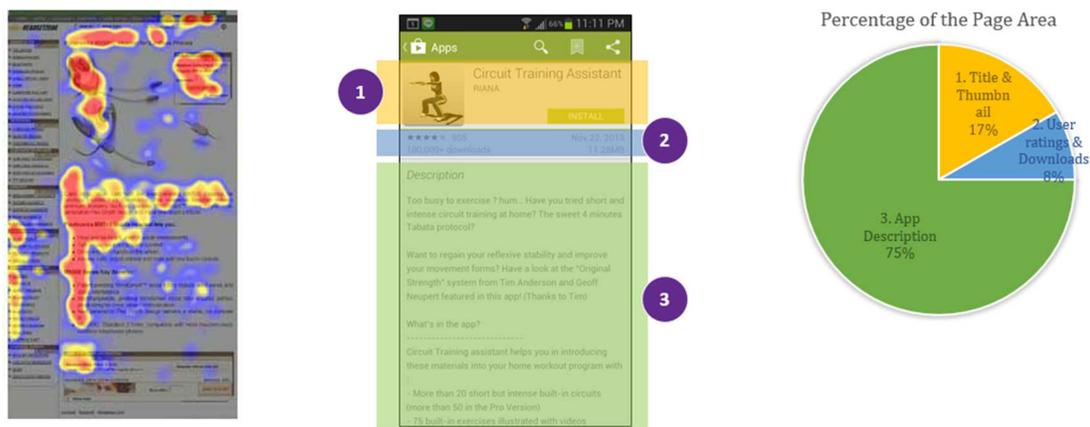
## Theoretical Background

### Mobile App Information

In general, mobile apps are sold via an app store, either website or mobile app store. An app’s information is presented on a page in the online store. This page is typically considered as the landing page that prompts users with a certain call-to-action features. In our case, it is the action to download the app. Landing pages must be well designed, visually appealing and easy to use since this is where the conversion occurs, which is the major goal of online commerce. In an app store, this page provides an app’s information which includes name, star ratings, description, price and developer information.

Although, among the app information presented on the app store page, the app’s star-ratings is treated as a relatively higher priority element than other components (Park et al. 2008), customers also observe the other product components displayed on the page. As found in the studies using eye-tracking approach, Nielsen and Loranger (2006) propose the F-shaped reading pattern that shows how people read information (especially product information) on the page (Figure 1). In addition, given that landing pages are critical to the success of business, mobile app descriptions typically occupy the largest space on the pages. Figure 1 presents a page real-estate analysis which demonstrates that mobile app descriptions take approximately 75% of the space on the entire page.

These findings are supported by literature in the initial trust area (McKnight et al. 2001; McKnight et al. 1998) which proposes that people look for product attributes (or cues) to constitute evidence of trustworthiness. This is especially the case when the customers do not have prior knowledge with the products.



**Figure 1. The F-Shaped Reading Pattern (Nielsen and Loranger. 2006) and a Page Real-Estate Analysis.**

## **Mobile App Descriptions**

We consider a mobile app description as a product's attribute that is created to stimulate and motivate potential customers to download the app. According to Hassenzahl (2003), there are two related aspects of product attributes: (1) pragmatic product attributes and (2) hedonic product attributes.

With respect to the pragmatic aspect, a mobile app description can be used to convey relevant functionality of the app, e.g., the app is "easy to use", "supportive", "useful", or "controllable". Such a pragmatic product attribute is primarily instrumental and is used to fulfil externally given or internally generated behavioral goals. This is consistent with suggestions from studies in the marketing domain that consider a product description as an attribute that signals quality of the product that ultimately leads to customers' purchase intention (e.g., Chang et al. 1994; Lee 2002).

Regarding the hedonic aspect, Hassenzahl (2003) proposed that hedonic attributes of products can influence individuals' emotional states. The hedonic aspect of software products can be divided into providing stimulation, communicating identity, and provoking valued memories. All these three dimensions of the hedonic aspect are generally expressed in mobile app descriptions in order to stimulate interest, raise attention, and evoke positive emotions of the customers.

## **Emotions in Mobile App Descriptions and Customer Behavioral Response**

Studies in the human-computer interaction (HCI) discipline suggest that emotions are an integral part of user experience when interacting with a computer system (Éthier et al. 2008; van der Heijden 2004). Emotion refers to a psychological reaction to events relevant to the needs, goals, or concerns of an individual (Brave et al. 2003). Human emotion is comprised of physiological, affective, and cognitive components (Brave et al. 2003). In this paper, the interpretation of "emotional response" is often technically termed sentiment, the assigned emotional property of an attribute (description) of mobile apps.

The relationship between emotions and user response can be explained by the affect-as-information model (Schwarz et al. 1983). Schwarz and Clore (1983) argue that emotional states serve as information. For example, when people are asked to report how much they like a product, they may base their judgment on their feelings about the product instead of reviewing its specific features. More generally, individuals may make judgments of virtually any target by assessing their feelings at that time and using those feelings as a basis for their attitudes (Wyer Jr et al. 1999). The affect-as-information model has been widely used in designing persuasive communication (e.g., product description) to influence a recipient's attitudes in supporting the message's recommendation (Petty et al. 1986).

Consequently, the relationship between attitudes and behavioral response can be explained by theory of reasoned action (TRA) (Fishbein et al. 1975), which hypothesizes a relationship between attitude and behavioral intention. TRA posits that intention is determined by three factors—attitude, subjective norms, and perceived behavioral control. Therefore, in this study, we propose that a user's emotion evoked by emotion words in a mobile app's description influences the user's attitude toward the app, which further impacts his/her behavior, such as downloading the app.

## **Sentiment Analysis of Web Document**

According to Pang and Lee (2008), sentiment analysis refers to the use of data mining techniques to identify and extract subjective information in source material. As a promising research area, text sentiment analysis has been extensively studied to detect sentiment in web documents. For example, Guha et al. (2004) applied sentiment analysis to examine trust in blogs. Miao et al. (2010) performed opinion-mining on unstructured and semi-structured user reviews. Cao et al. (2011) investigated the helpfulness of user reviews. Hu et al. (2012) studied how sentiment analysis can be used to detect fraudulent reviews. Mohammad (2012) analyzed emotions in mail and books. This study aims to further extend the application of text sentiment analysis into the area of persuasive communication, primarily mobile app descriptions.

## Methodology

Our methodology is mainly composed of data cleansing, emotion extraction using sentiment analysis, and text mining.

### Dataset

In this study, the original dataset contains information of 188,389 Android mobile applications. The data were collected in November 2011 by Frank et al. (2012). As of October 2011, this dataset encompasses approximately 59% of the Android App Store, which contained 319,161 active applications (Horn 2011). The application data were crawled from the web version of the Android App Store. For each application, the dataset contains the permissions requested by the application, the price, the number of downloads, the average user rating, and a description. The dataset consists of 30 different app categories.

Before performing data analysis, data cleansing process was applied to the dataset. In this phase, noise data and irrelevant data were removed. Confounding factors were also eliminated to minimize the confounding effects. These factors include: (1) app language, (2) price, and (3) star-rating. In this study, our methodology are applicable with text written in English language so app descriptions written in other languages were removed from the dataset. With regards to price, studies in the marketing literature consider price as the key extrinsic quality signal that strongly influences consumer use (Jacoby et al., 1971; Zeithaml, 1988). However, our research primarily focuses on the influence of app descriptions on the number of download. Therefore, the effect of price were controlled by including only free apps (price = \$0.00) in the dataset. In addition, we also considered the effect of star-rating. According to Chevalier and Mayzlin (2006), while the star-ratings affect intention to download, the impact of negative star-ratings (1-star) was found to be much stronger than those of moderate or positive star-ratings (3-5 stars), e.g., users simply ignore products with 1-2 stars. These findings suggest that the effect of star rating should be considered – either presence or absence. In our study, although it is ideal to eliminate the effect of star-rating in the analysis, this is not applicable in our dataset since each app is associated with a star rating. Thus, we decided to minimize such effect by including only apps with moderate to high star-ratings (3.5-5 stars). Finally, the dataset contains 6,000 unique apps and is further used for the analysis.

### Emotion Lexicon

The NRC Emotion Lexicon version 0.92 was adopted in this study for sentiment analysis. This is a large word-emotion association lexicon created by Mohammed and Turner (2010). The lexicon has entries for about 14,200 word types. Each word is classified as it associate to one or more of the following emotion categories: positive sentiment, negative sentiment, anger, fear, joy, sadness, disgust, surprise, trust, and anticipation.

Given the mobile app dataset, sentiment analysis involves determining which of the words exist in the emotion lexicon and calculates ratios such as the number of words associate with a particular emotion to the total number of emotion words in the text. According to Mohammed (2012), this approach may not be reliable in determining if a particular sentence is expressing a certain emotion, but such supervised classifying is efficient and reliable in determining if a large piece of text has more emotional expressions compared to others in a dataset.

In our mobile app dataset, each app is classified into one of the sentiment category according to the spikes of emotions detected in the app description.

## Analysis and Results

### *Salient Emotions in Android App Descriptions*

In this section, we quantitatively compare the emotion words in the app descriptions. Across the 30 app categories, approximately 90% of all the apps in the dataset is categorized into the anticipation sentiment category. This result is not surprising due to the nature of app descriptions that are normally written in persuasive communication fashion. Anticipation was then removed from the analysis so that the effects of other emotions can be further investigated.

The apps in each category was then analyzed according to the 7 remaining emotions – anger, disgust, sadness, surprise, joy, trust, and fear. As shown in Table 1, the emotion words that most frequently found include anger (1,667 apps), joy (1,986 apps), and trust (1,411 apps).

CategoryGroup	EmotionGroup							Total
	Anger	Disgust	Fear	Joy	Sad	Surprise	Trust	
Arcade & Action	322	15	54	94	1	11	18	515
Books & Reference	19	1	14	36	1	6	46	123
Brain & Puzzle	131	4	26	68	9	8	32	278
Business	16	0	4	12	0	1	13	46
Cards & Casino	62	3	1	18	1	0	15	100
Casual	148	13	33	100	7	12	9	322
Comics	14	2	0	10	0	2	14	42
Communication	35	7	21	74	1	4	76	218
Education	8	0	4	38	2	4	20	76
Entertainment	194	19	70	266	3	17	99	668
Finance	17	4	4	12	0	0	109	146
Health & Fitness	36	6	10	55	2	2	23	134
Libraries & Demo	9	5	4	17	0	3	12	50
Lifestyle	45	3	24	129	1	2	79	283
Media & Video	47	5	14	102	0	4	44	216
Medical	1	0	1	5	1	1	4	13
Music & Audio	36	1	6	142	0	1	18	204
News & Magazines	14	2	27	75	1	15	68	202
Personalization	72	17	54	191	6	6	87	433
Photography	17	4	4	38	0	1	10	74
Productivity	44	10	23	68	2	1	58	206
Racing	43	2	8	14	3	4	2	76
Shopping	17	6	12	58	1	3	25	122
Social	22	17	8	100	2	1	38	188
Sports	30	7	15	15	0	4	98	169
Sports Games	47	1	9	15	0	0	10	82
Tools	178	19	94	156	9	4	231	691
Transportation	7	2	5	9	0	1	27	51
Travel & Local	28	4	20	59	0	10	85	206
Weather	8	1	5	10	0	1	41	66
Total	1667	180	574	1986	53	129	1411	6000

**Table 1. Salient Emotions in Each App Category**

### ***Analysis of the Relationship between Emotion and Sentiment Polarity***

According to the previous analysis, as one of the negative polarity sentiments, anger appears to be the salient emotion in several app categories, e.g., Arcade & Action, Sport Games, or Cards & Casino. Therefore, further analysis was conducted to evaluate the relationship between anger and sentiment polarity. The results from Pearson's correlation indicate a strong relationship between anger and positive sentiment ( $p < 0.05$ ). These results contradict the findings from previous studies that generally consider anger as a negative sentiment (e.g., Tiedens 2001). Therefore, we suggest that the polarity of an emotion may be impacted by context, which, in this case, is the app category. Hence, future studies that investigate emotions should take context into account and carefully review the concepts of emotions before conducting the investigation.

### ***Influence of Emotions on App Download Counts***

The salient set of emotions (anger, joy, and trust) was then considered for further analysis. Given that the dataset comprises of 30 categories as shown in Table 1, a series of regression analysis was conducted to examine the impact of the salient emotion in each app category on its number of download. Before performing the regression analysis, outliers in the dataset were removed and the dependent variable

(number of download) was transformed so that it is normally distributed. Eventually, the dataset contains a total of 5,614 apps.

In general, the results exhibit significant relationship between the salient emotion and the number of downloads ( $p < 0.05$ ) in each app category. For example, as shown in Tables 2 and 3, trust was found to have a significant relationship with the number of downloads for apps in Tools and Utilities category ( $p = 0.038$ ), whereas joy was found to have a significant relationship with the number of downloads for apps in Entertainment category ( $p = 0.006$ ).

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-6341.003	31455.804		-.202	.840
	anger	-589.515	6919.614	-.006	-.085	.932
	joy	-3451.276	6942.782	-.035	-.497	.620
	trust	20614.062	9881.340	.143	2.086	.038

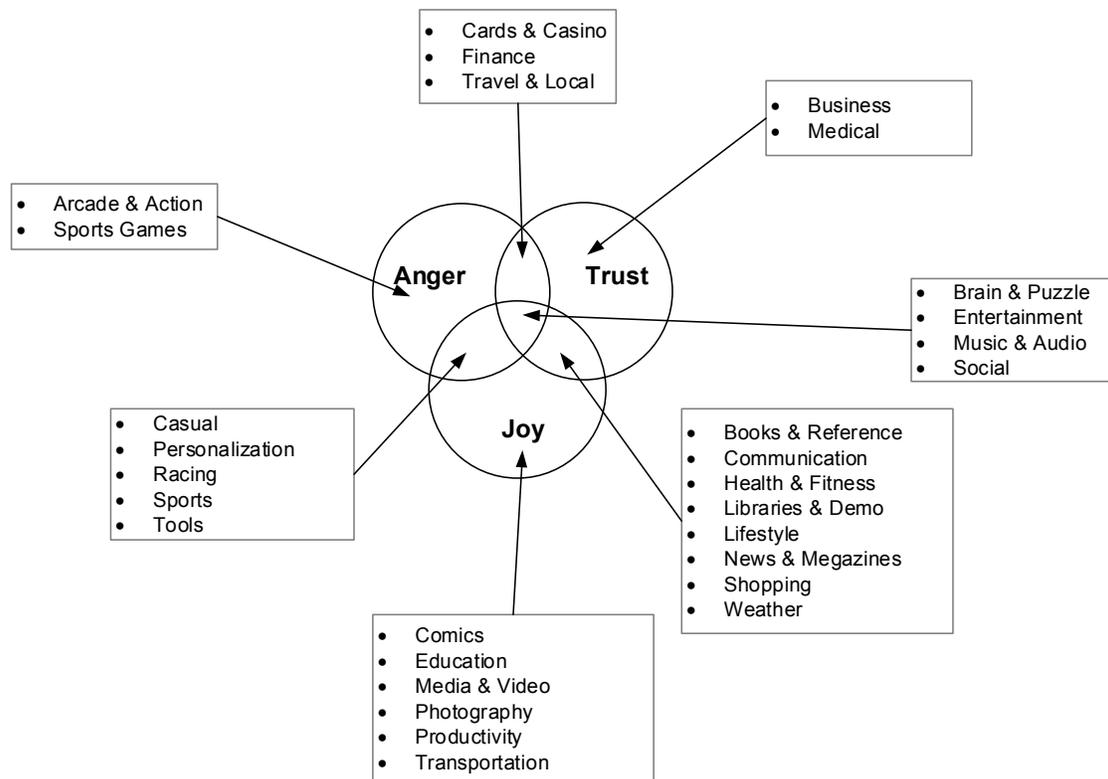
**Table 2. Significant Relationship between Trust and the Number of Downloads in Tools and Utilities Category**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-79942.286	42062.746		-1.901	.060
	anger	4769.070	9115.079	.050	.523	.602
	joy	40280.986	14385.888	.259	2.800	.006
	trust	11800.397	9437.769	.118	1.250	.214

**Table 3. Significant Relationship between Joy and the Number of Downloads in Entertainment Category**

***Salient Emotions in Top Downloaded Apps***

The results from the previous section indicate that there may be more than one emotions in a certain app categories relevant to the number of downloads. Therefore, further analysis was conducted to examine such effects of emotions for a more comprehensive understanding. Specifically, the focus was placed on the top downloaded apps in each app category. The apps in the dataset were ranked based on the number of downloads in each category, and then was analyzed by the stem-and-leaf plot. The results are presented in a Venn diagram of the three emotions (Figure 2). For example, all three combined emotions including anger, trust, and joy can potentially influence the number of download for apps in the Brain & Puzzle, Entertainment, Music & Audio, and Social categories.



**Figure 2. Salient Emotions in the Top 10% Downloaded Apps**

## Conclusion, Contributions and Future Research

In this study, a large word-emotion association lexicon was used to analyze and identify emotions in mobile app descriptions. This aspect of the online content generated from software vendors has been overlooked in the literature. Controlling for confounding factors (e.g., price and star ratings), the results exhibit a strong relationship between emotions and the number of an app downloaded. The causal effect of this relationship can be potentially explained by the affect-as-information model and theory of reasoned action.

Three most salient emotions – anger, joy, and trust - were identified from the dataset. We also found that the number of downloads can be potentially influenced by a combination of the salient emotions. We then compared these emotion combinations across all the 30 app categories in the dataset.

The contributions of the research is both theoretical and practical. The major theoretical contribute of this study is that we provide empirical evidence that emotions expressed in the persuasive communication have significant relationships with customer response. In addition, we also found that a negative emotion, anger, can potentially create positive sentiment, which depends on the context (app category). These findings suggest that future studies should take the context into account when emotions are investigated. Furthermore, the concepts and definitions of emotions must be carefully reviewed prior conducting future investigation to ensure both content and construct validities of the findings.

With regard to practical contributions, our study provides important managerial implications that are of interest to mobile app vendors, especially for newly established or unknown apps. Our findings suggest that app descriptions should be written with the emphasis on corresponding emotion words in each app

category. For example, words related to trust should be used frequently in writing descriptions for apps in the Business and Medical categories, whereas words related to joy should be used in descriptions for apps in the Education and Productivity categories.

Further work can be done based on this research in several areas. First, future study can apply sentiment analysis in searching snippets from multiple text in social media communication that have strong emotion word densities, e.g., Twitter feeds, Facebook posts, or text messages. Second, sentiment analysis can be applied to evaluate user emotions in a different context, both content generated by users (e.g., user reviews) and content generated by vendors (e.g., advertising text and webpage content). Third, sentiment analysis can be used to detect emotion in real-time communication between user and computer, and consequently, can be used to tailor specific service or experience that appropriate to the user's emotional states.

## REFERENCES

- Brave, S., and Nass, C. 2003. "Emotion in Human-Computer Interaction," in *The human-computer interaction handbook*, A. J. Julie and S. Andrew (eds.), L. Erlbaum Associates Inc., pp. 81-96.
- Cao, Q., Duan, W., and Gan, Q. 2011. "Exploring Determinants of Voting for the "Helpfulness" of Online User Reviews: A Text Mining Approach," *Decision Support Systems* (50:2), p 511.
- Chang, T. Z., and Wildt, A. 1994. "Price, Product Information, and Purchase Intention: An Empirical Study," *Journal of the Academy of Marketing Science* (22:1), pp 16-27.
- Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345-354.
- Éthier, J., Hadaya, P., Talbot, J., and Cadieux, J. 2008. "Interface Design and Emotions Experienced on B2C Web Sites: Empirical Testing of a Research Model," *Computers in Human Behavior* (24:6), pp 2771-2791.
- Fishbein, M., and Ajzen, I. 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, (Addison-Wesley Pub. Co. (Reading, Mass.))
- Frank, M., Ben, D., Felt, A. P., and Song, D. Year. "Mining Permission Request Patterns from Android and Facebook Applications," Data Mining (ICDM), 2012 IEEE 12th International Conference on 2012, pp. 870-875.
- Guha, R., Kumar, R., Raghavan, P., and Tomkins, A. 2004. "Propagation of Trust and Distrust," in *Proceedings of the 13th international conference on World Wide Web*, ACM: New York, NY, USA, pp. 403-412.
- Hassenzahl, M. 2003. "The Thing and I: Understanding the Relationship between User and Product," in *Funology: From Usability to Enjoyment*, M. Blythe, C. Overbeeke, A. F. Monk and P. C. Wright (eds.), Kluwer Academic Publishers: Norwell, MA.
- Horn, L. 2011. "Report: Android Market Reaches 500,000 Apps." Retrived from <http://www.pcmag.com/article2/0,2817,2395188,00.asp>.
- Hu, N., Bose, I., Koh, N. S., and Liu, L. 2012. "Manipulation of Online Reviews: An Analysis of Ratings, Readability, and Sentiments," *Decision Support Systems* (52:3), pp 674-684.
- Jacoby, J., Olson, J. C., and Haddock, R. A. 1971. "Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality," *Journal of Applied Psychology* (55:6), pp 570-579.
- Lee, P. M. 2002. "Behavioral Model of Online Purchasers in E-Commerce Environment," *Electronic Commerce Research* (2:1-2), pp 75.
- McKnight, D. H., and Chervany, N. L. 2001. "What Trust Means in E-commerce Customer Relationships: An Interdisciplinary Conceptual Typology," *International Journal of Electronic Commerce* (6:2), pp 35-59.
- McKnight, D. H., Cummings, L. L., and Chervany, N. L. 1998. "Initial Trust Formation in New Organizational Relationships," *The Academy of Management Review* (23:3), pp 473-490.
- Miao, Q., Li, Q., and Zeng, D. 2010. "Fine-grained Opinion Mining by Integrating Multiple Review Sources," *Journal of the American Society for Information Science and Technology* (61:11), pp 2288-2299.
- Mohammad, S. M. 2012. "From Once Upon a Time to Happily Ever After: Tracking emotions in mail and books," *Decision Support Systems* (53:4) 11, pp 730-741.

- Mohammad, S. M., and Turney, P. D. 2010. "Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon," in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Association for Computational Linguistics: Los Angeles, California, pp. 26-34.
- Nielsen, J., and Loranger, H. 2006. *Prioritizing Web Usability*, (New Riders Press: CA.
- Pang, B., and Lee, L. 2008. "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Informaiton Retrieval* (2:1-2), pp 1-135.
- Park, D. H., and Kim, S. 2008. "The Effects of Consumer Knowledge on Message Processing of Electronic Word-of-Mouth Via Online Consumer Reviews," *Electronic Commerce Research and Applications* (7:4), pp 399-410.
- Petty, R. E., and Cacioppo, J. T. 1986. "The Elaboration Likelihood Model of Persuasion," in *Advances in Experimental Social Psychology*, B. Leonard (ed.), Academic Press, pp. 123-205.
- Schwarz, N., and Clore, G. L. 1983. "Mood, Misattribution, and Judgments of Well-being: Informative and Directive Functions of Affective States," *Journal of Personality and Social Psychology* (45:3), pp 513-523.
- Tiedens, L. Z. 2001. "Anger and Advancement Versus Sadness and Subjugation: The effect of negative emotion expressions on social status conferral," (80), pp 86-94.
- van der Heijden, H. 2004. "User Acceptance of Hedonic Information Systems," *MIS Quarterly* (28:4), pp 695-704.
- Wyer Jr, R. S., Clore, G. L., and Isbell, L. M. 1999. "Affect and Information Processing," in *Advances in experimental social psychology, Vol. 31*, Academic Press: San Diego, CA, US, pp. 1-77.
- Zeithaml, V. A. 1988. "Consumer Perceptions of Price, Quality and Value: A Means-End Model and Synthesis of Evidence," *Journal of Marketing* (52:3), pp 2-22.