The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2633-0091.htm

Sentiment analysis of consumer reviews – a comparison of organic and regular food products usage

B. Rajeswari
Department of Management Studies, Pondicherry University, Pondicherry, India
S. Madhavan
Department of Management Studies, Manonmaniam Sundaranar University, Tirunelveli, India, and
Ramakrishnan Venkatesakumar and S. Riasudeen
Department of Management Studies, Pondicherry University, Pondicherry, India

Received 17 May 2020
Revised 6 June 2020
18 June 2020
Accepted 18 June 2020

Abstract

Purpose – This study aims to compare online review characteristics, review length and review sentiment score between "organic" and "regular" food products. In addition, variations in the consumer sentiment scores across the review lengths are studied.

Design/methodology/approach – This study fits into the descriptive research design. From Amazon's website, the consumer product reviews are scrapped. Using the text analytical package "sentiment" in R-Studio, we computed the sentiment scores and counted the number of words in each review. The mean sentiment scores and mean review length are compared for regular and organic products using one-way ANOVA. Sentiment score variation across review length and product class is studied through factorial ANOVA. Sample reviews of ghee and honey are used to test the hypotheses.

Findings – The review length shows a significant difference between the regular and organic products. The mean number of words in the regular products reviews is significantly lower than the mean number of words in the organic product reviews. The regular products' mean sentiment score is significantly lower than the mean sentiment score of organic products. The mean sentiment scores are not consistent between ghee and honey. Sentiment scores are better for organic honey and regular ghee products. For regular ghee products, longer reviews result in lower sentiment scores. For regular and organic versions of honey, longer reviews are associated with better sentiment scores.

Research limitations/implications – This study did not include the helpfulness of a review and the demographic data of the reviewers.

Practical implications – Sentiment scores' variations across the regular and organic and product categories such as ghee and honey give a comprehensive feedback to the firms. It also indirectly communicates a brand's evaluation by the consumers and the performance feedback for an upward extension like the organic category.

Social implications – Studies on organic category give feedback for environment-friendly products and consumer attitude shift towards safer products.

© B. Rajeswari, S. Madhavan, Ramakrishnan Venkatesakumar and S. Riasudeen. Published in *Rajagiri Management Journal*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at http://creativecommons.org/licences/by/4.0/legalcode

Rajagiri Management Journal
Vol. 14 No. 2, 2020
pp. 155-167
Emerald Publishing Limited
p-ISSN: 0972-9968
e-ISSN: 2633-0091
DOI 10.1108/RAMJ-05-2020-0022

Originality/value – Very limited studies have reported the upward line extensions. The authors study the upward line extension organic and associated sentiment scores variation. The role of review length and its systematic influence on the sentiment scores, variations in the review due to the product nature (organic/regular) are unique contributions of this study.

Keywords Organic, eWOM, Sentiment analysis, Grocery, Review length

Paper type Research paper

Organic consumer products are relatively a new category to the Indian context. Organic products are gaining their market shares in beverages, honey and food ingredients like ghee, dal and spices. These products are promoted with the catchy organic label, showing faster consumer acceptance. A joint report by the Association of Indian Chambers of Commerce (ASSOCHAM) and Ernst and Young have estimated that the organic market will have a cumulative aggregate growth rate (CAGR) of 25% and reach INR100bn by this year 2020. Many companies position organic products in the premium category and price them higher than the regular products.

The academic and marketing researcher address various changes in green and organic product purchase and consumption behaviour. A study result shows that the attitude is a lesser significant influencing factor on the environment than the knowledge component (Martin and Simintiras, 1995). The green consciousness, growing concern about a healthy lifestyle, self-realisation are the drivers of success of the organic products (Chan, 2001). In addition, apprehension for animal welfare, willingness to support local economy, nutritional value and ecological consciousness are encouraging organic product consumption (Laroche *et al.*, 2001). Premium price, poor distribution and low awareness about the certifying agencies by the consumers and poor marketing efforts are few concerns for the marketing of organic products.

Green or organic product consumers have different motives and consumption behaviour. The firms should also check the feasibility of entering the green market segments and serve them effectively (Sharma *et al.*, 2013). While promoting organic products, the marketers should emphasise environmental aspects and environment-friendly characteristics of these products (Eneizan *et al.*, 2019).

The surge in the use of smartphone and online platforms to share the views has brought out the challenges in managing the electronic word-of-mouth communications of consumers. Studies on social media behaviour of consumers prove that there is a shift in the information search behaviour and purchase process (Sturiale and Scuderi, 2013). During the purchase stage, people see comments of others on social media and decide their choices (Naz, 2015). A brand's participation and feedback on the social media or product review sites will have both positive and negative effect on the purchase decision (Jalilvand *et al.*, 2011).

Studies on online reviews are bringing out various dimensions of consumer behaviour and the relationship between variables. The length of a review and valence of the textual component affect the ratings given by the consumers (Yoon *et al.*, 2019). Helpfulness of the reviews vary with search and experience goods (Baek *et al.*, 2012).

Rather than attributes, specific attitudes and motives drive the organic product purchase and consumption behaviour. Such organic product use would result in different consumer reactions while posting online reviews. *Therefore, we expect that message characteristics of organic products will be different from that of regular products*. Specifically, we expect the sentiment scores and review lengths for organic products will be different from reviews of regular products.

In this study, we collected product reviews from the online retailer Amazon website. The consumers consider Amazon website as a trusted and credible source of information.

Among the product categories, we chose experience goods, and sample reviews are collected for ghee and honey. From the organic and regular versions of the product reviews, we have computed the sentiment scores and review length for further analyses. The study result provides insights for the organic product marketers, to manage the online customer reviews effectively.

2. Review of earlier studies and hypotheses

Usage of micro-blog sites, product review sites, and social media platforms gives a plethora of opportunities for the consumers to express their views, feedback and comments about the products. Managing this electronic word-of-mouth (eWOM) communication has become an important strategy for many marketers. Sales value is the key reason for the firms to keep an eye on eWOM (Arbelles *et al.*, 2020). To retain the brand trust, providing timely feedback on the online platforms is important for the firms (Bhandari and Rodgers, 2018).

Despite their popularity, social media sites are vulnerable to source credibility and quality. However, the blogs have the ability to normalise the counterarguments created for the products (Pant *et al.*, 2014). Comparison of social media and shopping site prove that eWOM information on shopping site is trusted more than social media information. The shopping sites are rated higher on content quality and source credibility (Erkan and Evans, 2018). These online reviews and product information are helpful to reduce consumer uncertainty during the purchase process (Sharma *et al.*, 2011) and capable of creating the patronising effect on retailers as well as nullifying negative information about the retailers (Chatterjee, 2001).

Marketing literature classifies products as search and experience goods (Sharma *et al.*, 2011). Usage of online information sources and other consumer reviews are varying across experience goods and search products. It is further moderated by the usefulness of the information sources (Bei *et al.*, 2004). No research studies address the variations in the online reviews within a category of goods.

Growing text analytics applications are broadening the understanding of the online reviews through new metrics and help in gaining new insights on the attitude of the consumers. One such metrics is sentiment analysis. Recent literature documented what is consumer sentiment (Bali *et al.*, 2016). Sentiment is a feeling, consciousness or thought about something. Researchers are addressing consumer reactions for products and services from the online reviews, feedbacks and blogging sites (Lin *et al.*, 2017). Text analytical tools like word clouds (Miley and Read, 2011), word storms (Castellà and Sutton, 2014) word extracting algorithms (Barth *et al.*, 2014), methodology to analyse sentiment scores (Lakshmi *et al.*, 2017) and emojis analysis framework (Li *et al.*, 2018) help the researchers to explore the online consumer reviews.

The review length, an important message characteristic, has been addressed in the recent studies. The research work on the review length establishes the relationship with the content quality (Hu *et al.*, 2008). Later, another research work on review length establishes its relationship with sentiment scores (Ghasemaghaei *et al.*, 2016). The review length is not an important factor in predicting the helpfulness of the review (Prakash *et al.*, 2016). This study addresses an important function of review length. It is proved that review length is a significant predictor of product performance.

Organic products are considered as the superior version of a regular product, and the consumers are driven by many unique motives and values to buy organic products. Hence, we expect organic products reviews may be longer than a regular product review and propose *H1* on review length and product versions:

H1. There is a significant difference in the mean number of words in the online reviews for regular and organic version of products.

To decide the purchase, many customers read at least ten product reviews (Hendrawan *et al.*, 2017). The sentiment scores for the customer reviews are used in many applications. Document-level sentiment scores are computed to understand the tone of the author (Feldman, 2013). Consumer reviews and sentiments in the reviews are influencing the purchase decision. Sentiment scores are also useful to identify positive and negative attributes of the products (Kau *et al.*, 2019). Thus, sentiment scores associated with the review is a proxy for the product performance or consumer satisfaction with the product.

The organic products have many positive features over the regular products. The organic product purchase is the result of deep thought process, willingness to pay higher prices and closer evaluation of various product attributes. Thus, we expect organic products reviews will have better sentiment than its *regular* version and propose the following hypothesis, *H2*:

H2. There is a significant difference in the mean sentiment score of online reviews between regular and organic versions of products.

The research works have established a relationship between various message characteristics with review length. Review length is related with content quality (Hu *et al.*, 2008), star rating (Baek *et al.*, 2012) and product performance (Prakash *et al.*, 2016). Review length is capable of affecting the helpfulness of the review (Prakash *et al.*, 2016) and sentiments in the messages (Ghasemaghaei *et al.*, 2016). Thus, we expect the review length will affect the sentiment scores.

A study on organic product proves that organic product consumption is a reflection of the values like environment consciousness, risk free products and willingness to pay a premium (Laroche *et al.*, 2001). When the consumers of organic products give reviews, it would be different from the other consumers who buy the regular versions. These two consumer segments will differ in terms of the product evaluation and the post-consumption behaviour. Sentiment score, being a proxy for consumer satisfaction (Al-Otaibi *et al.*, 2018), expect to differ between regular and organic product users.

Organic category being positioned as a superior quality version in a premium price range and the consumers expect to have better sentiment in the review messages than the messages from the consumers of the regular products. By combining these arguments on review length and product category, we expect that sentiment scores shall vary with review length and product versions. Thus, we propose the *H3*:

H3. There is a significant difference in the mean sentiment score across the product class and length of the review.

3. Research methodology

3.1 Research design

This research work compares the variations in the online consumers' reviews between the regular and organic products, on the review lengths and the sentiment scores of the reviews. Ghee and honey classified as experience category goods in the marketing literature and pantry category in Amazon are considered for the study. Both the products have niche variation – *organic*, apart from the regular offerings. These organic extensions are unique and mostly premium priced in the Indian market. Many regional or local brands and few national or MNCs brands characterise the competition. From the consumer online reviews,

we have collected the review text, given by the reviewers. For each review, we counted the review length and computed sentiment scores. Therefore, the study, with specific hypotheses and identified set of variables, fits into a *descriptive study* (Cooper and Schindler, 2002).

3.2 Sampling size and design

Indian arm of Amazon's (www.amazon.in) operation provides free samples up to 30–40% of the reviews available or to a maximum of 3,000 reviews. Applying purposive sampling method, sample reviews are collected through Google Chrome Scrapper's Amazon plugin.

Table 1 summarises the detail of samples reviews for ghee and honey and breakup across regular and organic products. The brand label carries the information of *Regular* or *Organic*. This classification is used as another study variable.

3.3 Study variables

Table 2 contains the brands and classification details across the product category chosen for analysis. Google Chrome Plugin has pre-specified fields like ID/Name/Text/Date/Header/rating. We used the review text only in this research. Once, the pre-processing is completed, the comments are simple *plain* English. Through another plugin "sentimentr", for each review, we have computed the sentiment scores and counted review length. These values are stored in the respective row and added back to the original data set. Further, the review length is divided into three categories: *Less than 5 words*, *6–10 words* and *More than 10 words*. Hence, we used *sentiment score*, *product category* (ghee/honey) and *product class* (regular/organic) and *review length* (three categories) as variables in the study.

Product		Product classification		
		Regular	Organic	Total reviews
Ghee	Count	3,684	1,387	5,071
	% of sample	72.6%	27.4%	
Honey	Count	2,749	1,961	4,710
	% of sample	58.4%	41.6%	
Total	Count	6,433	3,348	9,781
	% of sample	65.8%	34.2%	

Table 1. Classification of sample reviews

Product	Classification	Brands
Ghee	Regular	Nandini Pure Ghee, Nestle Everyday Ghee, Patanjali Cow' Ghee, GRB
Ghee	Organic	AMOREARTH, Brij Gwala, Brijmohan, Gau Organic, Gits Pure, Indicow Organic Dairy, Isha Vasyam A2 Ghee, Shree Radhey Gir Cow ghee, Trihoot Cow Ghee, Umanac Organic, Vanalaya Organic
Honey	Regular	Al Qusai, Dabur Honey, DYU Honey, Forever Living Products, Honey and Spice Wild Honey, Zandu Pure Honey
Honey	Organic	Bare Elixir Organic, DADEV Organic, First Bud Organics, Hi Honey, INDIGENOUS HONEY, Organic India, VANTATTVA, Wnature Organic

Table 2. Brands used in the study

3.4 Pre-processing the data

We used R-Studio and plugins like *tm*, *wordcloud*, *wordcloud2* and the command "gsub" to pre-process the text data. In line with earlier studies, review texts are treated for data cleaning process (Al-Otaibi *et al.*, 2018; Gaikar and Marakarkandy, 2015). This process includes removal of punctuations/special characters/numbers/symbols to lower-casing the words, removal of stem words and blank spaces. Then for each review, sentiment score is computed by the package *sentimentr* (Table 3).

It is a well-known fact that www.amazon.in gives the flexibility of filling the customer reviews, where, all the fields are not mandatorily to be filled by the reviewers. Hence, in some places, the sample sizes would be varying from the total reviews collected. Moreover, sentiment scores are not computed for incomplete reviews, single-letter reviews, numbers/symbol/emoji alone reviews and reviews in other languages. This issue resulted in different sample sizes than the total sample size collected. In the next section, we summarise the analysis of the data using ANOVA and factorial ANOVA.

4. Results

4.1 Results on number of words in the review

To compare the mean review length between the regular and organic products, one-way ANOVA is used with review length as dependent variable and product class as a factor variable. The mean number of words in the reviews are significantly higher for *organic* than *regular* versions of ghee and honey. These results support ***H1***. Thus, it is evident that *organic* is a niche category and attracted more number of words in the reviews than the *regular* products (Table 4).

4.2 Results on sentiment scores in the review

To find the variations in the mean sentiment scores of the reviews, one-way ANOVA is performed with the sentiment scores as the dependent variable and product class as the factor variable. The results confirm that there is a significant difference in the mean sentiment scores of *regular* and *organic* versions of ghee and honey products, and these findings support *H2*. For ghee, the *regular* products get a better mean sentiment score than

Table 3. Sample size across review length

Review length	Ghee	Honey
Less than 5 words	3,286	2,568
6–10 words	838	771
More than 10 words	928	1,358
Total	*5,052*	*4,697*

Table 4. Variations in review length

Product		Sample size	Mean review length	SD	F-ratio (significance)
Ghee	Regular	3,667	6.1110	9.5993	*115.132*
	Organic	1,385	10.3105	17.8271	*(0.000)*
	Total	5,052	7.2623	12.5487	
Honey	Regular	2,737	6.5543	9.7230	*226.548*
	Organic	1,960	13.1740	19.9355	*(0.000)*
	Total	4,697	9.3166	15.2161	

the *organic* version, whereas the results are reversed for honey. For honey, the *organic* products get a better mean sentiment score than the *regular* products. Table 5 summarises the results related to sentiment scores for the *regular* and *organic* versions of ghee and honey.

4.3 Results related to sentiment scores across product and review length

To analyse the variations in consumers' sentiment scores, we used a 2 × 3 factorial ANOVA with sentiment scores as dependent variable and the products versions (two l*evels: regular and organic*), review lengths *(three levels: less than 5 words, 6 – 10 words and more than 10 words) as* factor variables. The hypothesis (*H3*) on the sentiment scores *vis-à-vis* review length and product version has been supported partially for ghee and fully for honey.

From the factorial ANOVA, no significant difference is found in the mean sentiment scores of regular and organic ghee products *(F = 0.671, significance = 0.413)*, but mean sentiment scores are varying significantly across the groups based on review lengths *(F = 5.679, significance = 0.003)*. Significant interaction effect confirms that mean sentiment scores are varying across product class as well as review lengths (F = 26.762, significance = 0.000) (Tables 6 and 7).

The results for ghee products prove that shorter reviews carry better sentiment scores than longer reviews. However, the interaction effect suggests that, for *regular* products, shorter reviews possess better sentiment scores, whereas for *organic* products, longer reviews possess better sentiment scores.

Further from the factorial ANOVA results for honey, significant difference is found in the mean sentiment scores of *regular* and *organic* honey products *(F = 134.839, significance = 0.000)*. In addition, the mean sentiment scores are varying significantly across the groups based

Product		Sample size	Mean sentiment score	SD	F-ratio (significance)
Ghee	Regular	3,684	0.4074	0.3931	*9.346*
	Organic	1,387	0.3685	0.4306	*(0.002)*
	Total	5,071	0.3968	0.4040	
Honey	Regular	2,749	0.5132	0.4289	*273.393*
	Organic	1,961	0.7390	0.5048	*(0.000)*
	Total	4,710	0.6072	0.4752	

Table 5. Variations in sentiment scores

Products	Source of variation for sentiment score	F-ratio (significance)	Hypothesis result
Ghee	Product class	0.671(0.413)	*H3* is partially supported for ghee
	Review word group	*5.679(0.003)*	
	Product class * review word group	*26.762(0.000)*	
Honey	Product class	*134.839(0.000)*	*H3* is fully supported for honey
	Review word group	*158.525(0.000)*	
	Product class * review word group	*20.199(0.000)*	

Table 6. Summary of tests of between-subjects effects

Table 7.
Sentiment score variations for ghee

Product classification	Review length	Mean sentiment score	SD	No. of reviews
Regular	Less than 5 words	0.4503	0.3685	2508
	6–10 words	0.3305	0.4105	607
	More than 10 words	0.3098	0.4477	552
	Total	*0.4093*	0.3931	3667
Organic	Less than 5 words	0.3457	0.4112	778
	6–10 words	0.3549	0.4271	231
	More than 10 words	0.4260	0.4652	376
	Total	*0.3691*	0.4306	1385
Total	Less than 5 words	*0.4255*	0.3818	3286
	6–10 words	*0.3372*	0.4150	838
	More than 10 words	*0.3569*	0.4582	928
	Total	0.3983	0.4041	5052

on review lengths *(F = 158.525, significance = 0.000)*. The interaction effect is also significant, which suggests that the mean sentiment scores are varying across product class as well as review lengths (F = 20.199, significance = 0.000) (Table 8).

The results related to honey prove that shorter reviews possess lower sentiment scores than longer reviews. Sentiment scores are lower for *regular* honey and higher for *organic* honey products. A significant interaction effect confirms that for *regular* as well as *organic* honey products, shorter reviews possess lower sentiment scores and longer reviews possess better sentiment scores, and the sentiment scores are consistently higher for organic than regular products.

5. Discussion

In this research work, we address why some reviews are longer than other reviews. We have found out that reviews are longer for the organic ghee and honey. A study on green and organic product purchase intentions proves that green product awareness and brand image are positively related to the purchase intentions (Yang, 2017). By drawing insight from the elaboration likelihood model, we conclude that the consumers of the experience goods accumulate the product- and the usage-related information. Thus, the consumers' awareness about the product is higher, and they have a better experience with the organic products. These experiences and the augmented organic product knowledge resulted in longer

Table 8.
Sentiment score variations for honey

Classification	Review length	Mean	SD	No. of reviews
Regular	Less than 5 words	*0.4803*	0.3859	1837
	6–10 words	*0.4911*	0.4581	408
	More than 10 words	*0.6665*	0.5140	492
	Total	0.5154	0.4285	2737
Organic	Less than 5 words	*0.5439*	0.4019	731
	6–10 words	*0.6968*	0.4689	363
	More than 10 words	*0.9222*	0.5310	866
	Total	0.7394	0.5047	1960
Total	Less than 5 words	0.4984	0.3915	2568
	6–10 words	0.5880	0.4742	771
	More than 10 words	0.8296	0.5390	1358
	Total	0.6089	0.47480	4697

sentences to narrate the opinion (Nel and Boshoff, 2019). The marketers of the organic products shall consider the longer reviews as cues of the consumer acceptance and performance of the organic products closer to their expectations.

We sought an explanation for some online reviews contain a better sentiment than other reviews. We have found out that the mean score is lower for the regular version than the organic version. Marketing efforts like reaching out the consumers through SMS and e-mails improve the awareness level of organic products (Sturiale and Scuderi, 2013). A study also ascertains that the consumers are behaving responsible manner with increased concern for the environment (Fuiyeng and Yazdanifard, 2015). Consumer attitude is a significant factor for choosing the organic category (Kathuria and Gill, 2013), and there is a gradual shift in the consumerism and consumption of green products. From the study results, we conclude that the consumer awareness and justification for choosing the organic product is complemented by the better product performance, and this resulted in higher sentiment scores for the organic versions than the regular versions.

Moreover, we have found out that the sentiment scores for ghee is lower than the regular products, and longer reviews associated with lower sentiment scores. Ghee, an ingredient and experience good, lends itself for direct evaluation of very few intrinsic cues. Hence, the consumers felt the lower performance of the product against the expectations. On the other hand, honey being an experience good often consume directly, and such incidents are limited for ghee. Thus, the consumers could evaluate the performances of intrinsic cues directly in honey and articulated through the longer sentences and the longer reviews possess better sentiment than the shorter reviews.

6. Managerial implications

Social media, blogs, review sites and product sites facilitate consumers' feedback on products. The channels are capable influencing upon the consumer decision with a minimal effort (Subramani and Rajagopalan, 2003). "Word-of-mouse" is also getting equal importance as word-of-mouth among the habitual internet users. Marketers need to recognise the role of the internet in the consumer decision process and develop strategies exclusively for online platforms. In this study, we address two important online review characteristics: *review length* and *sentiment scores*.

6.1 Review length

Depending upon the product type, a reviewer's information diagnosticity is affected by the review length (Mudambi and Schuff, 2010). A study result on the review length is also reported that the longer reviews are often associated with extremely positive or positive reflections of consumers (Filieri, 2014). In this study, we have found out that the mean review length of organic versions of ghee and honey is longer than the regular products. This is a welcoming situation for the organic product marketers. Consumers spent their time, effort to register the positive feedback. Firms may identify such reviews and push these reviews to a special class of reviews as "top rated review" *or* "green review" category. This will encourage the customers to provide feedbacks with articulation on product attributes and patronising behaviour for organic versions.

6.2 Sentiment scores

Sentiment analysis becomes a hot area in decision-making process (Hussein, 2018). It is proved that star rating and review messages are not congruent with each other always (Baek *et al.*, 2012). To resolve this partial utility of star ratings, the marketers may develop procedures to show sentiment scores also in the review panel. This will help in breaking the

stereotyping error, believing organic version is always superior or regular product is inferior to organic version. At present, the product sites do not provide the sentiment scores for the product/brand. With the growing computing power and text mining applications, it would be apt to provide sentiment scores, which would be different from the star ratings given by the consumers.

7. Conclusion

Online consumer studies, being an emerging domain of research, establish relationships between message characteristics and consumer characteristics. Satisfactory performance of organic products and success of organic as a category in the Indian markets are established from the message characteristics like review length and sentiment scores.

The study findings are in parallel with the theoretical grounding on the elaboration likelihood model (ELM). Longer reviews associated with organic products are clear reflections of elaboration likelihood state of consumers, who think about issue-relevant information and his/her ability, motivation to elaborate the message (Nel and Boshoff, 2019). Thus, supports from the consumer satisfaction theory and ELM, the study findings conclude that performance of organic products is up to the expectations of consumers and consumer's preparedness to engage in positive eWOM.

The current findings are in line with earlier works, where it is proved that longer reviews are often associated with better product performances (Ghasemaghaei *et al.*, 2018), and the sentiment scores may be used as a proxy to measure the satisfaction (Al-Otaibi *et al.*, 2018).

8. Limitations and future research directions

The helpfulness of a review has its own magnitude of affecting the consumer eWOM. Because of the scope and objectives of the study, helpfulness is not considered in the present study. There are a growing number of machine learning algorithms on the text mining from the open sources to proprietary-based ones. These algorithms have their own initial and boundary settings. The results expect to have minor variations depending upon the algorithms. In the research, we used traditional algorithms available in the R-Studio plugins.

From the marketing perspective, analysing the data by combining the "cue theory" or "information search behaviour" could provide further insights on the online consumer behaviour. Using the product involvement or consumer knowledge or attitude or mood of the consumers can be value addition from the reviews. Future research works shall consider the mood of consumers as a variable by analysing the ratio of positive to negative words in the reviews and develop models on the online consumer behaviour.

References

Al-Otaibi, S., Alnassar, A., Alshahrani, A., Al-Mubarak, A., Albugami, S., Almutiri, N. and Albugami, A. (2018), "Customer satisfaction measurement using sentiment analysis", *International Journal of Advanced Computer Science and Applications*, Vol. 9 No. 2, pp. 106-117.

Arbelles, K.D., Berry, P. and Theyyil, A. (2020), "Electronic word-of-Mouth marketing on amazon: exploring how and to what extent amazon reviews affect sales electronic word-of-Mouth marketing on amazon: exploring how and to what extent amazon reviews affect sales", *The McMaster Journal of Communication*, Vol. 12, pp. 50-79.

Baek, H., Ahn, J. and Choi, Y. (2012), "Helpfulness of online consumer reviews: readers' objectives and review cues", *International Journal of Electronic Commerce*, Vol. 17 No. 2, pp. 99-126, doi: 10.2753/JEC1086-4415170204.

Bali, A., Agarwal, P., Poddar, G., Harsole, D. and Zaman, N.M. (2016), "Consumer ' s sentiment analysis of popular phone brands and operating system preference", *International Journal of Computer Applications*, Vol. 155 No. 4, pp. 15-19.

Barth, L. Kobourov, S.G. and Pupyrev, S. (2014), "Experimental comparison of semantic word clouds. In experimental comparison of semantic word clouds", In: Gudmundsson J., Katajainen J. *(eds)* Experimental Algorithms. *SEA 2014. Lecture Notes in Computer Science* (pp. 1-17).

Bei, L.-T., Etta, Y., Chen, E. and Widdows, R. (2004), "Consumers ' online information search behavior and the phenomenon of search vs. Experience products", *Early Childhood Education Journal*, Vol. 25 No. 4, pp. 449-467.

Bhandari, M. and Rodgers, S. (2018), "What does the Brand say? Effects of brand feedback to negative eWOM on Brand trust and purchase intentions", *International Journal of Advertising*, Vol. 37 No. 1, pp. 125-141, doi: 10.1080/02650487.2017.1349030.

Castellà, Q. and Sutton, C. (2014), "Word storms: Multiples of word clouds for visual comparison of documents", *In Proceedings of the 23rd international conference on World Wide Web*, pp. 665-675.

Chan, R.Y.K. (2001), "Determinants of Chinese consumers ' green purchase behavior", *Psychology and Marketing*, Vol. 18 No. 4, pp. 389-413.

Chatterjee, P. (2001), "Online reviews: do consumers use them?", *In ACR 2001 Proceedings, eds. M. C. Gilly and J. Myers-Levy, Provo, UT: Association for Consumer Research*, pp. 129-134.

Cooper, D.R. and Schindler, P.S. (2002), *Business Research Methods*, Tata McGraw-Hill, New Delhi.

Eneizan, B., Mohamad Alhamad, A., Bin.Mat Junoh, M.Z. and Binti Tunku Ahmad, T.S. (2019), "Green marketing strategies: theoretical approach", *American Journal of Economics and Business Management*, Vol. 2 No. 2, pp. 77-94, doi: 10.31150/ajebm.vol2.iss2.69.

Erkan, I. and Evans, C. (2018), "Social media or shopping websites? The influence of eWOM on consumers ' online purchase intentions", *Journal of Marketing Communications*, Vol. 24 No. 6, pp. 617-632, doi: 10.1080/13527266.2016.1184706.

Feldman, R. (2013), "Techniques and applications for sentiment analysis", *Communications of the Acm*, Vol. 56 No. 4, pp. 82-89, doi: 10.1145/2436256.2436274.

Filieri, R. (2014), "What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM", *Journal of Business Research*, Vol. 68 No. 6, doi: 10.1016/j.jbusres.2014.11.006.

Fuiyeng, W. and Yazdanifard, R. (2015), "Rly Ea", *Global Journal of Management and Business Research*, Vol. 15 No. 5, pp. 17-23.

Gaikar, D. and Marakarkandy, B. (2015), "Product sales prediction based on sentiment analysis using twitter data", *International Journal of Computer Science and Information Technologies*, Vol. 6 No. 3, pp. 2303-2313.

Ghasemaghaei, M., Eslami, S.P., Deal, K. and Hassanein, K. (2016), "Reviews ' length and sentiment as correlates of online reviews ' ratings", *Internet Research*, Vol. 28 No. 3, pp. 1-34.

Ghasemaghaei, M., Eslami, S.P., Deal, K. and Hassanein, K. (2018), "Reviews' length and sentiment as correlates of online reviews' ratings", *Internet Research*, Vol. 28 No. 3, pp. 544-563, doi: 10.1108/IntR-12-2016-0394.

Hendrawan, R.A., Suryani, E. and Oktavia, R. (2017), "Evaluation of E-Commerce product reviews based on structural, metadata, and readability characteristics", *Procedia Computer Science*, Vol. 124, pp. 280-286, doi: 10.1016/j.procs.2017.12.157.

Hu, N., Lie, L. and Zhang, J.J. (2008), "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects", *Information Technology and Management*, Vol. 9 No. 3, pp. 201-214, doi: 10.1007/s10799-008-0041-2.

Hussein, D.M.E.D.M. (2018), "A survey on sentiment analysis challenges", *Journal of King Saud University – Engineering Sciences*, Vol. 30 No. 4, pp. 330-338, doi: 10.1016/j.jksues.2016.04.002.

Jalilvand, M.R., Esfahani, S.S. and Samiei, N. (2011), "Electronic word-of-mouth: challenges and opportunities", *Procedia Computer Science*, Vol. 3, pp. 42-46, doi: 10.1016/j.procs.2010.12.008.

Kathuria, L.M. and Gill, P. (2013), "Purchase of branded commodity food products: empirical evidence from India", *British Food Journal*, Vol. 115 No. 9, pp. 1255-1280, doi: 10.1108/BFJ-08-2011-0209.

Kau, E., Peral, J., Gil, D., Ferr, A., Sellers, R. and Mora, H. (2019), "Managing marketing Decision-Making with sentiment analysis: an evaluation of the main product features using text data mining", *Sustainabiliity*, Vol. 11, pp. 1-19.

Lakshmi, V., Harika, K., Bavishya, H. and Harsha, C.S. (2017), "sentiment analysis of twitter data", *International Research Journal of Engineering and Technology*, Vol. 4 No. 2, pp. 2224-2227.

Laroche, M., Bergeron, J. and Barbaro-Forleo, G. (2001), "Targeting consumers who are willing to pay more for environmentally friendly products", *Journal of Consumer Marketing*, Vol. 18 No. 6, pp. 503-520, doi: 10.1108/EUM0000000006155.

Li, M., Ch'ng, E., Chong, A.Y.L. and See, S. (2018), "Multi-class twitter sentiment classification with emojis", *Industrial Management and Data Systems*, Vol. 118 No. 9, pp. 1804-1820, doi: 10.1108/IMDS-12-2017-0582.

Lin, C., Liaw, S., Chen, C., Pai, M., Chen, Y., Lin, C. and Chen, C. (2017), "RESEARCH PAPER a computer-based approach for analyzing consumer demands in electronic word-of-mouth", *Electronic Markets*, Vol. 27 No. 3, pp. 225-242, doi: 10.1007/s12525-017-0262-5.

Martin, B. and Simintiras, A.C. (1995), "The impact of green product lines on the environment: does what they know affect how they feel?", *Marketing Intelligence and Planning*, Vol. 13 No. 4, pp. 16-23.

Miley, F. and Read, A. (2011), "Using word clouds to develop proactive learners", *Journal of the Scholarship of Teaching and Learning*, Vol. 11 No. 2, pp. 91-110.

Mudambi, S.M. and Schuff, D. (2010), "What makes a helpful online review? A study of customer reviews on amazon.com", *MIS Quarterly*, Vol. 34 No. 1, pp. 185-200, doi: 10.2307/20721420.

Naz, F. (2015), *Word of Mouth and Softdrink*, Vol. 4 No. 1, pp. 1-4.

Nel, J. and Boshoff, C. (2019), "Journal of retailing and consumer services online customers ' habit-inertia nexus as a conditional effect of mobile- service experience: a moderated-mediation and moderated serial-mediation investigation of mobile-service use resistance", *Journal of Retailing and Consumer Services*, Vol. 47, pp. 282-292, doi: 10.1016/j.jretconser.2018.12.003.

Pant, A., Hsieh, Y., Lee, M. and Shen, C. (2014), "Pointing out differences: Consumer evaluations of comparative information across product types", *Journal of Promotion Management*, Vol. 20 No. 3, pp. 390-409, doi: 10.1080/10496491.2014.908802.

Prakash, J., Irani, S., Rana, N.P., Dwivedi, Y.K., Saumya, S. and Kumar, P. (2016), "Predicting the "helpfulness" of online consumer reviews", *Journal of Business Research*, Vol. 70, doi: 10.1016/j.jbusres.2016.08.008, pp. 1-10.

Sharma, R.S., Morales-Arroyo, M. and Pandey, T. (2011), "The emergence of electronic word-of-mouth as a marketing channel for the digital marketplace", *Journal of Information, Information Technology, and Organizations*, Vol. 6, pp. 41-61, doi: 10.28945/1695.

Sharma, V., Sonwalkar, J. and Kapse, M. (2013), "Consumer purchase behaviour for green products", *International Journal of Economics and Business Administration*, Vol. I No. Issue 4, pp. 50-65, doi: 10.35808/ijeba/25.

Sturiale, L. and Scuderi, A. (2013), "Evaluation of social media actions for the agrifood system", *In 6th International Conference on Information and Communication Technologies in Agriculture, Food and Environment (HAICTA 2013)*, Elsevier B.V, Vol. 8, pp. 200-208, 10.1016/j.protcy.2013.11.028

Subramani, M.R. and Rajagopalan, B. (2003), "Knowledge-sharing and influence in online social networks via viral marketing", *Communications of the Acm*, Vol. 46 No. 12, pp. 300-307, doi: 10.1145/953460.953514.

Yang, Y.C. (2017), "Consumer behavior towards green products", *Journal of Economics, Business and Management*, Vol. 5 No. 4, pp. 160-167, doi: 10.18178/joebm.2017.5.4.505.

Yoon, Y., Kim, A.J., Kim, J. and Choi, J. (2019), "The effects of eWOM characteristics on consumer ratings: evidence from TripAdvisor.com", *International Journal of Advertising*, Vol. 38 No. 5, pp. 684-703, doi: 10.1080/02650487.2018.1541391.

Corresponding author
Ramakrishnan Venkatesakumar can be contacted at: venkatesh1316@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**