Qualitative Evidence for a Behavioral Extension of the Expectancy Valence Theory: The Netflix-Twitter-Yakult Case

Ashish Varma

Institute of Management Technology, Ghaziabad, India
Email: avarma@imt.edu

Abstract

The Netflix movie “To all the boys I’ve loved before” released in the USA on 17th Aug 2018 was linked to a substantial increase in sales of the probiotic drink Yakult. This study probes the phenomenon with an analysis of tweets for the first three weeks after the release of the movie and also assesses reasons for the phenomenon with the theoretical underpinning of the Expectancy Valence (EV) theory. The first section of the study was conducted by undertaking a probe on 1922 tweets made during 17th Aug 2018 (the release date of the movie in the USA) till 6th September 2018. The second part of the study was conducted by taking intensive in-depth interviews of the target segment after 6th September 2018. The study highlights the growing significance of social media in shaping the sentiments by influencing perception and opinion (valence), which eventually influences the sales of products (expectation). The study found out that the sharp increase in sales was due to positive behavioral sentiment created by the social media.

Keywords
Expectancy-Value Theory, Twitter, Qualitative Method

1. Introduction

The social media has gained popularity with the youth across the world. The developments in the telecommunication field and the information technology field have further advanced the use of social media for communication purposes. This study is built around the Japanese dairy producer Yakult Honsha Company which is engaged in production of health drink Yakult since 1935. Yakult is marketed in a distinctive bottle with a foil top. It is a fermented milk drink with
A strain of “Lactobacillus casei” which is considered good for overall health and well-being. Few of the key characters in the Netflix movie, “To all the boys I’ve loved before” (released on 17th August 2018 in the USA), are shown drinking “a yogurt smoothie”. As the youth have always been influenced by the movies, the product immediately generated a lot of excitement, acceptance and had an influence on the target youth audience which supposedly led to an increase in sales. Consequently, the stock price of Yakult also increased sharply since the release date of the movie. Thus, there is a need to study if social media such as Twitter can positively influence the sales of a product by creating a positive sentiment.

It is very interesting to note that the movie does not mention “Yakult” at all. It is well understood that not all movie fans use twitter and vice versa yet, this phenomenon was quite remarkable. It is believed that on Twitter, approximately 6000 tweets are made per second and it is estimated that about 500 million tweets are made per day. Still “Yakult” could stand-out in all the noise. Twitter allows maximum 140 characters for a single tweet and it is very popular as a micro-blogging site especially amongst the youth audience.

A unique feature of the study is the qualitative design which was undertaken since there were many perspectives of the phenomenon which were yet to be discovered. Since the aim was to explore whether the social media could positively impact the sales of a product, a grounded theory approach was the best methodological choice for theory generation and elaboration. Another contribution of the study is that it improves the present understanding of the EV theory in an emerging market context. The grounded theory process starts when the first set of data is received and the further collection of the data is also influenced by the analysis of the data of the first set. There is no fixed minimum sample size and rather the focus is on purposive sampling to focus on cases which are information rich with respect to the study. In this study “key informants” were intensively interviewed to gain insights into the phenomenon. The study found out that social media, in this case Tweets, reinforced the positive behavioral sentiment about the movie and the product and also lead to an increase in sales of the product.

The paper is structured as follows. After the introduction in Section 1, conceptual development is discussed in Section 2 which is followed by the qualitative analysis of tweets and Yakult’s stock price movement in Section 3. Section 4 discusses the data analysis of the in-depth interviews and the grounded theory development. Section 5 discusses the findings, limitations and the discussion which is followed by the conclusion in Section 6.

2. Conceptual Development

A “sentiment” had been defined as a positive or negative feeling (Go et al.) [1] which collectively impact the acceptance of products and services. In the past, scholars have studied blogs and social networks which lead to opinion mining [2]. Other scholars such as Yang et al. [3] used web pages as the data for their
A. Varma

study. The concept of using tweets as qualitative inputs for a grounded theory development remained largely unexplored.

The Efficient Market Hypothesis (EMH) suggests that the stock prices are largely fluctuating on the basis of news, current events [4] and thus an understanding of the collective sentiment through tweets would be useful to predict such movements [5]. Other prior works such as those of Asur and Huberman [6] had predicted box office collections for movies prior to the release based on Twitter analysis. Bollen [7] studied whether the public sentiment, as studied from twitter, was correlated to the Dow Jones Industrial Index. Bing et al. [8] studied tweets and predicted industry’s sector-based stock prices. Poddar et al. [9] studied the impact of location-based tweets. Hence, there was a promise in using tweets as raw data in the theory building exercise.

Sotiriadis and Van Zyl [10] highlighted that social media, which is based on word of mouth, was more credible than other forms of paid media. However, none of these prior works used a qualitative interview method post the initial social media-based data collection to go deep into the cause of the phenomenon and to develop a grounded theory of social media’s impact on sales. The present study fills this important gap. The fact that Twitter’s API is open to other applications and most of the data is public [9] further positions tweets as interesting input for this study.

The present study does not aim to perform a sentiment analysis and rather presumes the sentiment to be present based on primary observation of tweets and the resultant stock price movement of Yakult. The objective of the study is to de-mystify the sentiment, and do a qualitative probe leading to an extension of the Expectancy Valence (EV) theory. The study also borrows from the pragmatic constructivism school. The construct causality is central to this school. In the context of this study, construct causality is evidenced by the integration of facts, human behavior, possible outcomes, and making social media influence the target audience.

The twitter user profile is understood to be of younger age groups [9] and such age groups were understood to be more influenced by their own peer groups [11]. The starting premise of the study was to evaluate the influence of the movie on its audience and subsequently creating the positive sentiment on twitter which proved substantial on creating a positive demand for Yakult. The above discussions lead to the following hypothesis:

H1: Tweets influence the creation of a positive sentiment.
H2: A positive sentiment leads to a favorable sales trend.

The tweet data was collected and analyzed by using NVivo [12]. To obtain deeper insights regarding the findings of the tweet data, personal in-depth interviews were subsequently conducted on the target audience.

3. Qualitative Analysis of Tweets and Yakult Stock Price Movements

A qualitative research design was chosen since the aim of the study was to build
and elaborate a theory. In this study both “how” and “why” questions needed to be addressed and the “sense making” exercise was undertaken by the author. The author used query, visualisation techniques and memos in the first stage of the data analysis by using the software.

The author also undertook extensive coding exercise. The twitter data has many dimensions and thus NVivo [12] was used for coding purposes rather than doing the coding manually. NVivo [12] also assists in organizing, deep analysis for finding patterns in data and presentation of findings. Coding from “ten thousand feet view” with a broader picture of phenomenon led to some very interesting findings. A total of 1992 records (tweets) were studied by making 844 codes with references totaling to 8857. The data for 21 days, from 17th August 2018 to 6th September 2018 (both days included) were collected. The 17th August date was the movie release date in the USA. Although pre-release sentiments might have some impact, it was not considered for the purpose of this study. It is also to be noted that in any social media platform, 21 days is not a very short period of time as the topics of discussion changes very frequently on the social media. The data was carefully screened and all irrelevant tweets were eliminated to improve the quality of data used for the purposes of the study. An illustration of the code book is given in Figure 1.

The tweets were captured using “NCapture” Add-in without causing any obstruction and inconvenience to the tweeters. Search for hashtag, keyword, and username was undertaken. For instance, one of the keywords searches were made for “Yakult Sales Netflix” and the second search was made for “Yakult + To All the Boys I’ve Loved Before”.

Finally, two “NCapture” sets were made to collect the relevant tweets. Both the data sets were eventually merged to eliminate multiple and duplicate tweets. A purposive sampling process was undertaken which was followed by an assessment of the topics/themes discussed in the tweets. The data had to be carefully collected as the possibility of misspelling and slangs is high in tweets [1]. Similar items were clustered together using the “Cluster Analysis” method is shown in Figure 2.

The items clustered by coding similarity were calculated on the basis of Jaccard’s coefficient [13] as the similarity metric (Figure 3).
As per the terms and conditions of Twitter, only the derivations and conclusions from the data after processing the same were discussed in the paper. Therefore, no raw data or individual tweeter is discussed for the purposes of the study—no matter how prolific or significant they may be. This research was conducted purely for academic purposes to help theory building process. Again, as per Twitter terms of usage, only the public tweets are used for this study. This was manageable since most twitter accounts are public, and there was no need to purchase any data from twitter. The tweeter always has the right to withdraw the tweet at any point in time. However, a big point of contention amongst the scholars is whether the twitter data was primary data or secondary data. The author does not resolve this conflict as it’s beyond the scope and objectives of this study. The Bloomberg terminal was accessed to ascertain the trend in the stock price movement of Yakult (Figure 4). The stock price shows a rising trend during first 21 days of release of the movie.

![Items clustered by coding similarity.](image1)

**Figure 3.** Items clustered by coding similarity.

![Yakult stock price movement. Source: Bloomberg accessed on 11 Sep 2018.](image2)

**Figure 4.** Yakult stock price movement. Source: Bloomberg accessed on 11 Sep 2018.
4. Data Analysis of the In-Depth Interviews and the Grounded Theory Development

Multiple data sources are a must for such a study to ensure rigor in qualitative data analysis, and therefore apart from the tweets, personal in-depth interviews were also undertaken to build the grounded theory. The target respondents were between 18 to 25 years of age. The respondents were chosen to have a nearly equal representation of males and females in the sample interviewed. The author was open to the idea that the research propositions could also be refined/modified/extended above and beyond the expectancy valence theory (EV) as the study progressed. There was an intended suspension of judgment after the literature review so as not to prejudice the findings of the interviews with pre-decided notions of the researcher. Only the initial research questions were grounded in the existing literature. There were many follow-up questions that were asked to the respondents. The questions asked were predominantly open ended however a few close ended questions were also asked to ensure clear understanding of the responses. For instance, one question asked was "Please share an instance when you got influenced to try a product based on the media frenzy surrounding a movie?". The aim of this question was to understand and observe the respondent’s emotional intensity and experiences behind the purchase decision. The respondents were interviewed in their natural setting i.e. their colleges and institution of higher learning. Each interview lasted from 20 minutes to 30 minutes.

The interview method [14] which advocated “going native” and explore deeply and at the same time keeping the authors role as less as possible was initiated for the informant who were considered as the “knowledge agent”. However, the informant reported all what was observed/discussed. To the interview of the semi-structured interviews, complete anonymity was promised to get

Table 1. "Data Structure” of a select few demonstrative quotes from qualitative interviews leading to theory building.

<table>
<thead>
<tr>
<th>The 1st order Concepts (Informants words)</th>
<th>2nd Order themes which relate fully or partially to existing theory. (Authors Key Phrases as Researcher);</th>
<th>Aggregate Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent 1 (Female, 18, Bachelors degree student)</td>
<td>&quot;I found the novel to be so very cool, adorable, heartwarming and is one of my favorite stories&quot;. &quot;I have no idea about Yakult although I sort of remember Yogurt in the Netflix movie&quot;.</td>
<td>Emotional connect, Behavioral dimension</td>
</tr>
<tr>
<td>Respondent 2 (Female, 22, Business Executive)</td>
<td>“The movie was so perfect and so nicely adopted from the novel”</td>
<td>Positive re-enforcement</td>
</tr>
<tr>
<td>Respondent 3 (Male, 19, B.Tech student)</td>
<td>“I am sometimes swayed into trying some products and I do have some major likes, but it’s not very often”</td>
<td>Short term motivation to try a product</td>
</tr>
<tr>
<td>Respondent 4 (Female, 23, Entrepreneur)</td>
<td>&quot;If I like the movie, I can buy the novel, I would sometimes, but not always, try the products endorsed in the movie&quot;. “The movies do affect me, emotionally&quot;.</td>
<td>Behavioral tension, Human nature, Respondents narratives of past behavior</td>
</tr>
</tbody>
</table>

DOI: 10.4236/tel.2018.815212 3462 Theoretical Economics Letters
free and honest responses. The notes taken and observations made were shown back to the respondents and then transcribed. The interview protocol revision process [15] was also undertaken.

As suggested in the Gioia Methodology [16] the author also made a revision of the interview questions as per the progress made in the previous interviews and also as per the primary “Gestalt Analysis” i.e. the whole is more comprehensive and meaningful and so should be interpreted collectively. A coding exercise was undertaken to maintain the integrity of the 1st Order Terms which was followed by a comprehensive “compendium of 1st Order Terms”. After that the author converted the 1st Order into the 2nd Order (Theory Centric) themes. First a “bottom-up” inductive process was used to develop patterns in the data which was followed by a “top-down” deductive logic to integrate themes with the data captured. The final result of the process was the creation of data structure by identification of principles that are portable. Thus, for the final reporting purposes, the data structure creation [16] was attempted as under:

- The 1st order concepts (Informants words)
- The 2nd order theories (Authors Key Phrases as researcher)
- Aggregate Dimensions
- Finally, the second round of literature review was undertaken for confirming the “overall concept” and the “smaller and more specific construct” to conclude the research process.

5. Findings, Limitations and Discussion

Social media analytics involve collection of data from the social media to gain insights of users tastes and preferences. The research illustrates the use of what comes close to being called as “tweet related big data” for developing a primary and quick understanding of a phenomenon. It then extends that understanding with in-depth interviews for a theory development exercise which validates and extends the expectancy valence theory (EV) with the help of qualitative data.

The study had some limitations as well:

- The first limitation of the study is that data was collected for 21 days only and observation is based on short term trend.
- The second limitation of the study is that the data was collected for English language only.
- Finally, a challenge of microblogging based studies is the range of topics [17] and this study was no exception to this.

A word cloud was created to visually analyse the most prominent words of the tweets. As can be seen from Figure 5, the words Yakult, boys, loved, Netflix, Yog, can be very prominently seen from the word cloud. Although there are other prominent words as well, only these words are relevant for the purpose of this study.

The words Yakult, Netflix, sales, boys, loved are mentioned among the ten most frequently used words in the 1992 tweets (Figure 6). These words contribute
to the positive sentiment for Yakult and the increase in sales eventually giving support to both hypothesis 1 and hypothesis 2.

The word frequency (Table 2) clearly highlights that the 6-letter word “Yakult” was mentioned 6.3% weighted percentage with an actual count of 5239. This was the word with the highest word frequency among the 30 most frequently used words. The other prominently used words relevant for this study are yogurt (word count of 712; weighted percentage level of 0.86%), drink (word count of 598; weighted percentage 0.72%) shelves (word count of 365; weighted percentage level 0.44%), and fly (word count of 322; weighted percentage level 0.39%).

The coding by Twitter (Figure 7) highlights the presence of a couple of dominant tweeters. However, the individual tweeters identity, profile and objective of the tweets are beyond the scope of discussion in this study and thus were not probed further.

A detailed “geo assessment” of tweets was undertaken to ascertain the impact of the location-based tweets as it was relevant to know origin of the tweets. The map based on tweets (Figure 8) clearly shows that the tweets were a global phenomenon and were not restricted to a single city or country.
Figure 7. Coding by Twitter user on the basis of number of tweets.

Table 2. Word frequency of the top 30 words.

<table>
<thead>
<tr>
<th>Word</th>
<th>Length</th>
<th>Count</th>
<th>Weighted Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yakult</td>
<td>6</td>
<td>5239</td>
<td>6.30</td>
</tr>
<tr>
<td>sales</td>
<td>5</td>
<td>3584</td>
<td>4.31</td>
</tr>
<tr>
<td>loved</td>
<td>5</td>
<td>3203</td>
<td>3.85</td>
</tr>
<tr>
<td>boys</td>
<td>4</td>
<td>3173</td>
<td>3.82</td>
</tr>
<tr>
<td>since</td>
<td>5</td>
<td>3011</td>
<td>3.62</td>
</tr>
<tr>
<td>@netflix</td>
<td>8</td>
<td>2991</td>
<td>3.60</td>
</tr>
<tr>
<td>mentioned</td>
<td>9</td>
<td>2645</td>
<td>3.18</td>
</tr>
<tr>
<td>never</td>
<td>5</td>
<td>2617</td>
<td>3.15</td>
</tr>
<tr>
<td>@bloombergquint</td>
<td>15</td>
<td>2612</td>
<td>3.14</td>
</tr>
<tr>
<td>soaring</td>
<td>7</td>
<td>2605</td>
<td>3.13</td>
</tr>
<tr>
<td>lara</td>
<td>4</td>
<td>2594</td>
<td>3.12</td>
</tr>
<tr>
<td>jean</td>
<td>4</td>
<td>2593</td>
<td>3.12</td>
</tr>
<tr>
<td>yog</td>
<td>3</td>
<td>2536</td>
<td>3.05</td>
</tr>
<tr>
<td>https</td>
<td>5</td>
<td>2398</td>
<td>2.89</td>
</tr>
<tr>
<td>netflix</td>
<td>7</td>
<td>2135</td>
<td>2.57</td>
</tr>
<tr>
<td>yogurt</td>
<td>6</td>
<td>712</td>
<td>0.86</td>
</tr>
<tr>
<td>film</td>
<td>4</td>
<td>646</td>
<td>0.78</td>
</tr>
<tr>
<td>booming</td>
<td>7</td>
<td>600</td>
<td>0.72</td>
</tr>
<tr>
<td>drink</td>
<td>5</td>
<td>598</td>
<td>0.72</td>
</tr>
<tr>
<td>helps</td>
<td>5</td>
<td>405</td>
<td>0.49</td>
</tr>
<tr>
<td>teen</td>
<td>4</td>
<td>397</td>
<td>0.48</td>
</tr>
<tr>
<td>shelves</td>
<td>7</td>
<td>365</td>
<td>0.44</td>
</tr>
<tr>
<td>comedy</td>
<td>6</td>
<td>359</td>
<td>0.43</td>
</tr>
<tr>
<td>fly</td>
<td>3</td>
<td>322</td>
<td>0.39</td>
</tr>
<tr>
<td>seen</td>
<td>4</td>
<td>303</td>
<td>0.36</td>
</tr>
<tr>
<td>movie</td>
<td>5</td>
<td>286</td>
<td>0.34</td>
</tr>
<tr>
<td>look</td>
<td>4</td>
<td>272</td>
<td>0.33</td>
</tr>
<tr>
<td>gracias</td>
<td>7</td>
<td>256</td>
<td>0.31</td>
</tr>
<tr>
<td>came</td>
<td>4</td>
<td>236</td>
<td>0.28</td>
</tr>
</tbody>
</table>
A cluster analysis based on word similarity (Figure 9) shows that the words used in the tweets are quite intermixed in the different tweets. As per Figure 9, the words make a complex web, deeply interwoven, and thus the words are used repeatedly in the tweets in different combinations.

As advocated by Gioia et al. [16] there should be a deliberate delay in using codes that are based on existing theory (in this case the EV theory) to the last stage of data analysis. They suggest that the information obtained from the respondents should not be diluted with the burden of confirming with existing theories. This way the ideas are grounded in data rather than prior research studies. The author also used the existing theory in the last stage of data analysis in conformity with the “Gioia Methodology” which suggests value, in semi-ignorance of existing literature.

As per the EV theory, individuals make behavioral choices on the basis of anticipated outcomes (valences) and their futuristic estimations (expectancies) of attaining those outcomes [18]. This was evident from both the twitter data as well as the interview data. The respondents wanted to identify with the characters of the movie and the drink helped them achieve this goal. Thus, companies can position their products where the target audience wishes to identify itself with. As per the EV theory, higher motivation is associated with a higher perceived value for the rewards and the higher perceived probability that efforts will lead to rewards. As observed by the analysis of the tweets and the analysis of the interviews, the movie did motivate the target audience into using the product. Griffin and Harell [19] stated that additive form of expectancy model accurately predicts the level of motivational force acting on an individual. This was also the case here. From the data it can be concluded that the tweets impact the sales of products by initiating positive behavioral sentiments. As highlighted by Feather and O Brien [20] in the context of expectancy valance analysis, there was also a requirement to take into consideration the dynamic processes that influence behaviour over time. However, this can only be explored by future researchers.
Figure 9. Cluster on the basis of word similarity.

Parker and Roffey [21] suggested that grounded theory aims to integrate the researchers understanding for an explanation of the structures and processes observed thereby interpreting the data instead of simply reporting it. In the context of the present study the “dynamic relationship” of cause and effect, between positive sentiment reinforced by tweets and sales increase can be expected in future as well. To this extent, this theory of tweets-based sales increase is generalizable and predictable. There is also a significant potential for discovery about this theory in other contexts which can be taken up by future researchers.

6. Conclusions

Twitter is a popular website for microblogging where users create status messages (called “tweets”) [1]. The tweets are however with a maximum limit of 140 characters of text. The Netflix-Twitter-Yakult triad is a good case illustration of the power of the social media in positively influencing the intended audience. This finding has significant ramifications for strategic planning purposes as well. For the short run, the social media buy in, is a good strategy. It also appears that since firms innovate to create competitive advantage and managers take risks even in highly controlled environments to drive such innovation [22], the innovative influence of social media extends to strategic planning domain for long run decisions as well. Social media is an impactful and innovative method to positively influence the target audience especially ever since mobile technologies have started being used for both social and professional networking [23] by individuals and professionals alike. The advancements in mobile technology have made the use of social media very convenient.

The consumption of social media generated information and the resultant actions, is not limited to the youth only. For instance, management accountants have started appreciating big data [24] for decision making and entrepreneurs have adopted new technologies such as mobile banking [25] for their financial needs. Thus there seems to be a deep integration of communication technology,
mobile technology and information technology which is continuously re-shaping the world we live in. The study highlights the growing significance of social media in shaping the sentiments by influencing the perception and the opinion (valence), which eventually influences the sales of products (expectation). The study found out that the reasons for such immediate sales linked stock price increase were positive behavioral sentiments created by the social media.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References


