Vaporous Marketing: Uncovering Pervasive Electronic Cigarette Advertisements on Twitter

Eric M. Clark,1,2,3,4,5,6 Chris A. Jones,1,6,7,8 Jake Ryland Williams,1,2,3,4,5 Allison N. Kurti,1,8
Michell Craig Nortotsky,1,6 Christopher M. Danforth,1,2,3,4,5 and Peter Sheridan Dodds1,2,3,4,5

1University of Vermont
2Department of Mathematics & Statistics
3Vermont Complex Systems Center
4Vermont Advanced Computing Core
5Computational Story Lab
6Department of Surgery
7Global Health Economics Unit of the Vermont Center for Clinical and Translational Science
8Vermont Center for Behavior and Health
(Dated: August 11, 2015)

Background
Twitter has become the “wild-west” of marketing and promotional strategies for advertisement agencies. Electronic cigarettes have been heavily marketed across Twitter feeds, offering discounts, “kid-friendly” flavors, algorithmically generated false testimonials, and free samples.

Methods
All electronic cigarette keyword related tweets from a 10% sample of Twitter spanning January 2012 through December 2014 (approximately 850,000 total tweets) were identified and categorized as Automated or Organic by combining a keyword classification and a machine trained Human Detection algorithm. A sentiment analysis using Hedonometrics was performed on Organic tweets to quantify the change in consumer sentiments over time. Commercialized tweets were topically categorized with key phrasal pattern matching.

Results
The overwhelming majority (80%) of tweets were classified as automated or promotional in nature. The majority of these tweets were coded as commercialized (83.65% in 2013), up to 33% of which offered discounts or free samples and appeared on over a billion twitter feeds as impressions. The positivity of Organic (human) classified tweets has decreased over time (5.84 in 2013 to 5.77 in 2014) due to a relative increase in the negative words ‘ban’, ‘tobacco’, ‘doesn’t’, ‘drug’, ‘against’, ‘poison’, ‘tax’ and a relative decrease in the positive words like ‘haha’, ‘good’, ‘cool’. Automated tweets are more positive than organic (6.17 versus 5.84) due to a relative increase in the marketing words like ‘best’, ‘win’, ‘buy’, ‘sale’, ‘health’, ‘discount’ and a relative decrease in negative words like ‘bad’, ‘hate’, ‘stupid’, ‘don’t’.

Conclusions
Due to the youth presence on Twitter and the clinical uncertainty of the long term health implications of electronic cigarette consumption, the protection of public health warrants scrutiny and potential regulation of social media marketing.

PACS numbers:
ions regarding e-cigarettes over time. Individuals post tweets which are short text based messages restricted to 140 characters. Using data mining techniques, roughly 850,000 tweets containing mentions of e-cigarettes were collected from a 10% sample of Twitter’s garden hose feed spanning from January 2012 though December 2014. This analysis extends a preliminary study [24] which analyzed all e-cigarette related tweets spanning May through June 2012.

As Twitter has become a mainstream social media outlet, it has become increasingly enticing for third parties to gamify the system by creating self-tweeting automated software to send messages to organic (human) accounts as a means for personal gain and for influence manipulation [25]. We recently introduced a classification algorithm that is based upon three linguistic attributes of an individual’s tweets [26]. The algorithm analyzes the average hyperlink (URL) count per tweet, the average pairwise dissimilarity between an individual’s tweets, and the unique word introduction decay rate of an individual’s tweets.

All tweets mentioning e-cigarettes were categorized using a two-tier classification process. Tweets containing an abundance of marketing slang (‘free trial’, ‘starter kit’, ‘coupon’) are immediately categorized as automated. All of the tweets from individuals that have mentioned an e-cigarette keyword are collected in order to classify the remaining tweets per individual as either organic or automated. The machine learning classifier was trained on the natural linguistic cues from human accounts to identify promotional and SPAM entities by exclusion.

The emotionally charged words that contribute to the positivity of various subsets of tweets from each category were quantitatively measured using hedonometrics [27] [28]. Outliers in both the positivity and frequency time-series distributions correspond to political debates regarding the regulation of e-cigarettes. Recent studies [29–33] report an alarmingly rapid increase in the youth awareness and consumption of electronic cigarettes; a Michigan study found that the use of e-cigarettes surpass tobacco cigarettes among teens [34]. The CDC reports that “the number of never-smoking youth increased three-fold from approximately 79,000 in 2011 to 263,000 in 2013” [35]. During this time-period there has also been a substantial (256%) increase in youth exposure to electronic cigarette television marketing campaigns [36]. Due to the high youth presence on Twitter [37] as well as the clinical uncertainty regarding the risks associated with e-cigarettes, understanding the effect of promotionally marketing vaporization products across social media should be immediately relevant to public health and policy makers.

Materials and Methods

Data Collection

An exhaustive search from the 10% “garden hose” random sample of Twitter spanning 2012 through 2014 yielded approximately 850,000 tweets mentioning a keyword related to electronic cigarettes including: e(-)cig, e(-)cigarette, electronic cigarette, etc. All tweets were tokenized by removing punctuation and performing a case insensitive pattern match for keywords. Using timezone meta-data the tweets were converted into their local post time, in order for a more accurate ordinal sentiment analysis. The language, reported by Twitter, and user features were also collected and analyzed.

Automation Classification

As reported in [24] there is a high prevalence of automation among e-cigarette related tweets. Many of these messages were promotional in nature, offering discounted or free samples or advertising specific electronic cigarette paraphernalia (see Table 3). A human detection algorithm defined and tested in [26] was implemented to classify accounts as either automated or organic (human in nature). All tweets from each individual appearing in our dataset were collected for the classifier. For each individual, the average URL count, average tweet dissimilarity, and word introduction decay rate were calculated for the individuals with at least 25 sampled tweets.

The majority (94%) of commercial e-cigarette tweets collected by [24] contain a hyperlink (URL). The average URL count per tweet has been demonstrated to be a strong feature for detecting robotic accounts [38–40]. Many algorithmically generated tweets contain similar structures with minor character replacements and long chains of common substrings, as opposed to Organic content. The Pairwise Tweet Dissimilarity of tweets $t_i, t_j$ from a particular individual was estimated by subtracting the length (number of characters) of the longest common subsequence, $|LCS(t_i, t_j)|$ from the length of both tweets, $|t_i| + |t_j|$ and normalizing by the total length of both tweets:

$$D(t_i, t_j) = \frac{|t_i| + |t_j| - 2 \cdot |LCS(t_i, t_j)|}{|t_i| + |t_j|}.$$  

For example, given the two tweets:

$(t_1, t_2) = (I\ love\ tweeting, \ I\ love\ spamming)$. Then $|t_1| = 16, |t_2| = 15, LCS(t_1, t_2) = |I\ love\ | = 7$ (including whitespace) and we calculate the pairwise tweet dissimilarity as:

$$D(t_1, t_2) = \frac{16 + 15 - 2 \cdot 7}{16 + 15} = \frac{17}{31}.$$
The average tweet dissimilarity of the individual was then estimated by finding the arithmetic mean of each individual’s calculated pairwise tweet dissimilarity. Since automated and promotional accounts have a structured and limited vocabulary, the unique word introduction decay rate introduced in [41] serves as another useful attribute to detect automated accounts. Using these attributes, the calibrated human detection algorithm, tested in [26], detected over 90% of automated accounts from a mixed 1000 user sample with less than a 5% false positive rate.

The Human Detection Algorithm was calibrated for a range of tweet sample sizes from hand classified Organic accounts. Ordinal samples of collected tweets from each account were binned into partitions of 25 ranging from 25 to a maximum of 500 tweets. Table 1 below lists the number of automated and organic classified accounts per year. Individuals with less than 25 sampled tweets were not classified with the detection algorithm.

To benchmark the accuracy of the detection algorithm on this sample of tweets, a random sample of 500 accounts algorithmically classified as automatons and 500 classified as Organic were hand coded to gauge the accuracy of the detection algorithm. The feature set of each sampled individual is plotted in three dimensions. The traced box indicate the organic feature cutoff. True Positives (red) are correctly identified automatons, True Negatives (green) are correctly identified Humans, False Negatives (blue) are automatons classified as humans and False Positives (orange) are humans classified as automatons.

The Categorization by Topics

Tweets with at least 3 advertising jargon references (e.g. coupon, starter kit, free trial) were immediately classified as automated. All posts from users with at least 10 marketing classified tweets were also flagged as automated. As noted in [24], some Organic users could retweet promotional content for rewards (e.g. winning free samples or discounts). All of these tweets were still classified as automated, but the user was not flagged as such. The remaining tweets were classified as either automated or organic by the human detection algorithm. Posts from users who had an insufficient number of sampled tweets (< 25) to algorithmically classify and who hadn’t posted commercial content were classified as Organic. Due to the high prevalence of hyperlinks included in tweets from promotional accounts, Tweets with URLs whose user had insufficient tweets to classify algorithmically were discarded (3.85% total tweets). A final list with each tweet classification coding is created by merging the commercial keyword classification with the results from the Human Detection Algorithm.

Results and Discussion

The number of automated, and in particular promotional, tweets vastly overwhelm (80.7%) the organic (see Figure 2). The identified automated accounts tweet e-cigarette content with much higher frequency than the Organic users. The average number of automated tweets

### TABLE I: Human Detection Twitter Account Classification

<table>
<thead>
<tr>
<th>Year</th>
<th>Automated</th>
<th>Organic</th>
<th>Unclassified*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>12,715</td>
<td>12,052</td>
<td>19,512</td>
</tr>
<tr>
<td>2013</td>
<td>64,874</td>
<td>59,376</td>
<td>120,142</td>
</tr>
<tr>
<td>2014</td>
<td>54,033</td>
<td>63,289</td>
<td>48,528</td>
</tr>
</tbody>
</table>

*account had less than 25 sampled tweets
per user was 1.96 with a standard deviation of 35.06 and a max of 14,310. Average organic posts per user were 1.44 with a standard deviation of 4.01 and max of 356 tweets. A total of 607,446 Automated Tweets provided a URL (92.09%).

Frequency WordClouds (see Figure 3) illustrate the most frequently used words by the Automated category. The size of the text reflects the ranked word frequencies. Marketing key words (Free Trial, Brand, Starter Kit, win, Sale) and brand names (V2, Apollo) are prevalent, illustrating commercial intent. Many automated tweets also refer to the health benefits of switching to electronic cigarettes (#EcigsSaveLives), even though they have not been officially approved as such by the Food and Drug Administration, \cite{42, 43}. See Table 3 for sub categorical counts of the automated tweets.

Tweet Sentiment Analysis

Hedonometrics are performed on the organic subset of electronic cigarette tweets to quantify the change in user sentiments over time. Using the happiness scores of English words from LabMT \cite{27}, along with its multi-language companion \cite{28} the average emotional rating of a corpus is calculated by tallying the appearance of words found in the intersection of the word-happiness distribution and a given corpus, in this case subsets of tweets. A weighted arithmetic mean of each word’s frequency, $f_{\text{word}}$, and corresponding happiness score, $h_{\text{word}}$ for each of the $N$ words in a text yields the average happiness score for the corpus, $\bar{h}_{\text{text}}$:

$$
\bar{h}_{\text{text}} = \frac{\sum_{w=1}^{N} f_{w} \cdot h_{w}}{\sum_{w=1}^{N} f_{w}}
$$

The average happiness of each word, $h_{\text{avg}}$ lies on a 9 point scale: 1 is extremely negative and 9 is extremely positive. Neutral words ($4 \leq h_{\text{avg}} \leq 6$), aka ‘stop words’, were removed from the analysis to bolster the emotional signal of each set of tweets.

Figure 4 shows that automated electronic cigarette tweets are using very positive language to promote their products. The average happiness of the Organic tweets are much more stable, and are becoming slightly more negative over time. Both distributions have a sudden drop in positivity during December 2013, around a debate regarding new e-cigarette legislation by the European Union. These tweets, labeled #EuEcigBan, are investigated separately in the next section. The words
that have the largest contributions to changes in sentiments are investigated with Word-shift graphs.

Word-shift graphs, introduced in [27], illustrate the words causing an emotional shift between two word frequency distributions. A reference period ($T_{ref}$), creates a basis of the emotional words being used to compare with another period, ($T_{comp}$). The top 50 words responsible for a happiness shift between the two periods are displayed, along with their contribution to shifting the average happiness of the tweet-set. The arrows ($\uparrow$, $\downarrow$) next to a word indicate an increase or decrease, respectively, of the word’s frequency during the comparison period with respect to the reference period. The addition and subtraction signs indicate if the word contributes positively or negatively, respectively, to the average happiness score.

In Figure 5, below, Word-shift graphs compare the change in Organic sentiments over time, as well as the difference in sentiments between automated and organic tweets. On the top, the 2013 Organic Tweet distribution is used as a reference to compare sentiments from 2014 Organic Tweets. December 2013 and January 2014 are removed to dampen the effect of tweets mentioning the #EUecigBan (see Figure S1). The average happiness score decreases from 5.84 in 2013 to 5.77 in 2014. This decrease in the average happiness score is due to a relative increase in the negative words ‘ban’, ‘tobacco’, ‘doesn’t’, ‘drug’, ‘against’, ‘poison’, ‘tax’; a relative decrease in the positive words ‘haha’, ‘good’, ‘cool’. Notably, there is also relatively less usage of the words ‘quit’, ‘addicted’, and an increase in ‘health’, ‘kids’, ‘juice’. On the bottom, Organic tweets from 2013 is the reference distribution to compare Automated tweets from the same year. Automated tweets are more positive (6.17-6.59 versus 5.84) due to a relative increase in the marketing words ‘best’, ‘win’, ‘buy’, ‘sale’, ‘health’, ‘discount’, etc and a relative decrease in the negative words ‘bad’, ‘hate’, ‘stupid’, ‘don’t’, among others. The words ‘free’ and ‘trial’ are excluded from the graph, since their high frequency and happiness scores distort the image ($h_{avg}$ increases from 6.17 to 6.59).

Sub-Categorical Tweet Topics

Pertinent topics related to e-cigarette marketing regulation include kid-friendly flavors, smoking cessation claims, and price reduction (including free trials, and starter kits). Keywords from each of these topics are used to sub-classify the automated tweet set per year, see Table 3 below. Purely commercial tweets were those with any marketing keywords including: ‘buy’, ‘save’, ‘coupon(s)’, ‘discount’, ‘price’, ‘cost’, ‘deal’, ‘promo’, ‘money’, ‘sale’, ‘purchase’, ‘offer’, ‘review’, ‘code’, ‘win(ner)’, ‘free’, ‘starter kit(s)’, ‘premium’. The URL from each tweet was also analyzed for promotional keywords. Any URL with at least three mentions of the above keywords was enough to classify the tweet as commercial.

When an individual on Twitter ‘follows’ another account, posts from these users appear on the ‘timeline’ of the individual. We quantify the social reach of each of these sub-categorical tweets by counting the total number of accounts’ ‘timelines’ who could have been exposed to the advertisement. To approximate this, we sum the number of followers from each individual’s tweets. The total number of impressions from the commercial category increases from 195.25 million to 951.03 million between 2013 to 2014, even though the total count has dropped from 283k to 149k. This implies that promotional accounts that are successful in deceiving Twitter’s SPAM detector may be gaining many more social links to broadcast their commercial context.

In order to gauge the accuracy of these sub-categorical tweet topics, 500 tweets were randomly sampled from each category and were evaluated separately by two people to determine the relevance of the tweet to its categorization. The evaluators had a high level of concordance (84.8%) and the discrepancies were resolved and merged into a final list. Sampled tweets were highly relevant per category, the percentage for each is given in Table 3 below.

Many automated tweets mentioned using electronic cigarettes as a cessation device, or as a safe alternative. Over 20,000 tweets were classified as cessation related, which potentially appeared on over 76.8 million individ-
FIG. 5: (Top) Organic Tweets from 2013 are the reference distribution to compare sentiments of Organic Tweets from 2014 where we see a negative shift in the calculated average word happiness. The computed average happiness \( \bar{h}_{avg} \) decreases from 5.82 to 5.77 due to both an increase in the negative words ‘tobacco’, ‘drug’, ‘ban’, ‘poison’, and a decrease in the positive words ‘love’, ‘like’, ‘haha’, ‘cool’ etc. (Bottom) Organic Tweets from 2013 are the reference distribution to compare Automated Tweets from 2013.

<table>
<thead>
<tr>
<th>TABLE III: Automated Tweet Subcategory Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subcategory</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>Commercial</td>
</tr>
<tr>
<td>Discount</td>
</tr>
<tr>
<td>Flavor</td>
</tr>
</tbody>
</table>

*Relevant percentage of 500 randomly sampled tweets

Electronic cigarettes have not been conclusively authorized as an effective cessation device, [11] has demonstrated the ineffectiveness of electronic cigarettes to suppress nicotine cravings. It is also notable that these affiliate marketing accounts are advertising electronic cigarettes as a completely safe alternative to analog tobacco use, contrary to recent studies [44–47]. Cessation tweets were tallied using the keywords ‘quit’, ‘quitting’, ‘stop smoking’, ‘smoke free’, ‘safe’, ‘safer’, ‘safest’. Many of the purely commercialized tweets mentioned discounts or even free samples. These Discount tweets were categorized with the keywords ‘free trial’, ‘coupon(s)’, ‘discount(s)’, ‘save’, ‘sale’, ‘free (e)lectronic (cig)arette’. Tweets advertising flavors were tallied using the keywords ‘flavor(s)’ and ‘flavour(s)’.

A noteworthy class of E-cigarette commercial-bots, are those that are masquerading as Organic users to spam pseudo-positive messages towards potential consumers. These “cyborgs”, as defined in [26, 38], spam a positive message regarding a personal experience. One class of these automatons are sending contrived testimonies that e-cigarettes have successfully allowed them to quit smoking cigarettes. These messages are very intentionally structured and tend to swap a few words to appear organic. These messages also target specific individuals as a more personal form of marketing. The general tweet structure from a sample cyborg marketing strategy is given below:

- @USER \{I,We\} \{tried,pursued\} to \{give up, quit\} smoking. Discovered BRAND electronic cigarettes and quit in \(#\) weeks. \{Marvelous,Amazing,Terrific\}! URL
- @USER It’s now really easy to \{quit,give up\} smoking (cigarettes). - these BRAND electronic cigarettes are lots of \{fun,pleasure\}! URL
- @USER electronic cigarettes can assist cigarette

A noteworthy class of E-cigarette commercial-bots, are those that are masquerading as Organic users to spam pseudo-positive messages towards potential consumers. These “cyborgs”, as defined in [26, 38], spam a positive message regarding a personal experience. One class of these automatons are sending contrived testimonies that e-cigarettes have successfully allowed them to quit smoking cigarettes. These messages are very intentionally structured and tend to swap a few words to appear organic. These messages also target specific individuals as a more personal form of marketing. The general tweet structure from a sample cyborg marketing strategy is given below:
smokers to quit, it’s well worth the cost URL

• @USER It’s {incredible, amazing} - the (really) {easy, painless} {answer, method} to quit cigarette smoking through BRAND electronic cigarettes URL

• I managed to quit smoking with these e-cigarettes, I highly recommend them: URL @USER

• @USER Its {amazing, extraordinary} - I (really) quit smoking after {#} yrs thanks to BRAND electronic cigarettes! URL

Using cyborgs to mimic Organic Users for marketing purposes should be analyzed heavily, to gauge their impact and effectiveness on consumers.

Conclusion

Our study has identified an abundance of automated, and in particular, promotional tweets, and consequent organic sentiments. The collected categorized tweet data from this analysis is available for follow-up analyses into e-cigarette social media marketing campaigns. Future work can perform a deeper analysis on the URL content, similar to [22], posted by promotional accounts to get a better sense of the smoking cessation, flavor mentions, and discount prevalence. We take care not to downplay the well recognized health benefits from smoking cessation including: decreased risk of coronary artery disease, cerebrovascular disease, peripheral vascular disease, decreased incidence of respiratory symptoms such as cough, wheezing, shortness of breath, decreased incidence of chronic obstructive pulmonary disease, and decreased risk of infertility in women of childbearing age [15] [18] [19]. The greatest concern of promotional e-cigarette marketing on Twitter is the risk of enticing younger generations who otherwise may never have commenced consuming nicotine. Due to the unknown but unignorable long-term adverse health effects of electronic cigarettes and the alarmingly increased youth consumption of these products, monitoring and potentially regulating social media commercialization of these products should be immediately relevant to public health and policy agendas.

Acknowledgements

The authors wish to acknowledge the Vermont Advanced Computing Core which provided High Performance Computing resources contributing to the research results. EMC was supported by the UVM Complex Systems Center, PSD was supported by NSF Career Award # 0846668. CJ, AK is supported in part by the National Institute of Health (NIH) Research wards R01DA014028 & R01HD075669, and by the Center of Biomedical Research Excellence Award P20GM103644 from the National Institute of General Medical Sciences.

* Electronic address: eclark@uvm.edu
[18] US Department of Health and Human Services and others, Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking
and Health 2 (2010).


[38] K. Zezima, Cigarettes without smoke, or regulation (2009), URL http://www.nytimes.com/2009/06/02/us/02cigarette.html?_r=2&.


Supplementary materials

European Union E-cigarette Ban Political Debate (#EUecigBan)

Each categorical time-series exhibits a severe negative trend occurring between December 2013 and January 2014. There is an inverse relationship with the average happiness scores during this time period. This was during the time that the EU was debating strict regulation and a possible ban on specific e-cigarette products [12]. Hashtags (#) allow users to categorize the content of their tweets. During this period, 13,227 sampled tweets were tagged with #EUecigBan. In Figure S1, a word shift graph (left) visualizes the sentiments from English Organic users using #EUecigBan versus the remaining Organic tweets from 2013. English Tweets tagged #EUecigBan are the comparison distribution in reference to all other tweets from 2013. Tweets containing #EUecigBan are on average much more negative (h_avg 5.81 versus 5.37) due to an increase in the negative words 'ban', 'stop', 'no', 'not', 'fight', 'against', 'disaster', 'death', 'corruption', 'tobacco', 'kills', etc. The positive words also disfavor the legislation, with the words 'save', 'millions', 'lives', 'support', 'healthy' occurring more frequently. English, French, and German tagged tweets were the most prevalent, and word clouds help visualize themes between language and user class (see Figure S1). This shows that Twitter sentiments can be useful in gauging public opinion toward regulation of electronic cigarettes. There is also a heavy automated tweet presence in each language with a similar attitude regarding the legislation, as depicted in the word clouds in Figure S1. Future work should also investigate if and how automated users can impact organic opinion on legislation.

![Figure S1 1](Left) Word shift graph comparing tweets tagged #EUecigBan against 2013 English Organic User Tweets (untagged). (top-right) The automated and Organic tagged tweet distributions are plotted. A histogram displays the counts per language and user class. (bottom-right) Word clouds compare ranked-word frequencies across language and user type.