The Chicken or the Egg: The Relationship between Traditional and Social Media Campaigns, and Evidence of Behavior Change

Sherry L. Emery, PhD
Change Institute
May 13th, 2014
Teamwork makes the dream work...
Roadmap: Overview and Highlights

• Background
  – How do we measure ‘exposure’ today? A New-ish Paradigm for Evaluating Media Effects
  – Collecting Social Media Data: Why How We Do It Matters

• Observing the conversation about electronic cigarettes on Twitter

• Observing tobacco control conversations on Twitter

• Other platforms and why they matter

• Punch Line and Directions for Future Research
Conceptual Framework Pre-2011

Demographic Controls

Contextual/Policy Controls

Message Exposure → Attitudes Beliefs → Outcomes

#HMCTalks
New Paradigm ➔ New Conceptual Framework

Demographic Controls

[Diagram showing gears labeled Exposure, Share, and Search, and an arrow connecting to Health Behavior]

Contextual/Policy Controls
Why Twitter is Great

- One of Fastest Growing SM Platforms
  - 23% of Internet users (up from 18% in 2013)
    - 27% Black
    - 37% 18-24
- Both Commercial and Individual Content
- Inherently a public platform
  - Spontaneous and unfiltered information
  - One in three Twitter users posted tweets about the content of television programs while viewing (Bauder, 2012).
- Fewer Technical Challenges
There are Lots of Ways to Get Social Media Data

- And the choice matters....
  - Streaming API
  - Search API
  - Firehose
  - Garden Hose

How you get your data defines what you can analyze
USE-CASE 1:

ELECTRONIC CIGARETTES ON TWITTER

#HMCTalks
Questions About E-Cigs Messages:

• How much electronic cigarette promotion is there on Twitter?

• How much ‘organic’ conversation about electronic cigarettes is there on Twitter?

• How does Twitter electronic cigarette messaging/promotion fit into the larger media landscape?
Dataset Compilation

1. Generate a broad pool of keywords
   - Expert input, focus groups, search traditional and alternative sources (e.g., urbandictionary.com)

2. Screen the keywords for relevance and frequency

3. Discard keywords that returned poor results (>70% irrelevant)

4. Add new keywords as necessary

5. Go back to Step 2
Keywords for E-Cigs

#HMCTalks
@hmcollab
Pilot: Two Months of E-Cigs on Twitter 2012

Data Collection
• May 1-June 30, 2012.

Human Coding and Machine Classification
• Human coders reviewed a sample of 5,000 Twitter messages, overlapping on 500 messages
• Used human codes to train naïve Bayes classifier
• Classifier was highly consistent with human codes (kappa ≈ 0.9)

Content
• Keyword rules used to assess fraction that mentioned health, safety, smoking cessation, pricing/discounts.
## Characteristics of Organic and Commercial E-cig Tweets

<table>
<thead>
<tr>
<th>Tweet Type</th>
<th>Total tweets</th>
<th>Total users</th>
<th>Average e-cig tweets per user</th>
<th>Average followers</th>
<th>Total reach (sum of followers of all tweets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial/advocacy</td>
<td>66,102</td>
<td>17,936</td>
<td>3.7</td>
<td>848.4</td>
<td>173,133,016</td>
</tr>
<tr>
<td>Organic</td>
<td>7,570</td>
<td>6,254</td>
<td>1.2</td>
<td>866.9</td>
<td>3,834,216</td>
</tr>
<tr>
<td>All</td>
<td>73,672</td>
<td>23,700</td>
<td>3.1</td>
<td>844.7</td>
<td>176,967,232</td>
</tr>
</tbody>
</table>
**Top 10 Most Active Twitter Accounts**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Total e-cig tweets</th>
<th>Percent of tweets</th>
<th>Cumulative percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>13198</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>User 2</td>
<td>3363</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>User 3</td>
<td>1754</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>User 4</td>
<td>1326</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>User 5</td>
<td>825</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>User 6</td>
<td>704</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>User 7</td>
<td>630</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>User 8</td>
<td>626</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>User 9</td>
<td>547</td>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td>User 10</td>
<td>515</td>
<td>1</td>
<td>32</td>
</tr>
</tbody>
</table>

@hmcollab  
#HMCTalks
Can We Do Better?
And What Happened in 2013?
One Year of E-Cigs on Twitter

Data Collection
• October 1, 2012 - September 30, 2013.

Human Coding and Machine Classification
• Human coders reviewed a sample of 5,000 Twitter accounts, overlapping on 500 accounts.

• Used human codes to train three types of machine classifiers to classify the full set of 1,248,198 accounts.

Content
• Keyword rules used to identify reference to health, safety, smoking cessation, pricing/discounts, specific brands, or hashtags.
An Aside on Data/Analysis Quality

The SVM algorithm performed the best compared to human coders and was used to classify the accounts.

Precision = .98
Recall = .95
F1 score = .96
Amount and Fraction of E-cig/vape Product Accounts (1.2 million accounts)
Amount and Fraction of E-cig/vape Tweets (approx 5 million)
## Content Characteristics

<table>
<thead>
<tr>
<th>Tweet Content</th>
<th>Organic</th>
<th>Commercial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Claims</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health/Safety</td>
<td>53,906 (5.7%)</td>
<td>52,079 (1.4%)</td>
</tr>
<tr>
<td>Cessation</td>
<td>49,889 (5.3%)</td>
<td>70,290 (1.9%)</td>
</tr>
<tr>
<td>Price/Discount</td>
<td>63,685 (6.8%)</td>
<td>1,753,889 (47.5%)</td>
</tr>
<tr>
<td>Include URLs</td>
<td>290,293 (30.9%)</td>
<td>3,597,940 (97.4%)</td>
</tr>
<tr>
<td>Re-tweet</td>
<td>243,079 (25.9%)</td>
<td>3,092,574 (83.7%)</td>
</tr>
</tbody>
</table>

- Commercial tweets linked to vendor sites
- Commercial accounts heavily promoted discount starter kits, encouraging product uptake
- Commercial accounts showed a marked reduction in health/safety and cessation claims between the 2-month study from 2012 and this study
Red = commercial
Blue = social media
Green = organic
Interpreting the Network Map

• Nearly 1 millions tweets linked to FreeEcigaretteStarterKits.com

• Organic links focused around other social media sites and news outlets

• Commercial accounts frequently linked to YouTube.com
Other Platforms
Monthly electronic cigarette/vaping videos: Jan 2007 - Dec 2014
Expenditures Don’t Capture Exposure

Quarterly promotional spending for e-cigarettes, 2010-2013

Source: Kantar Media 2015

#HMCTalks @hmcollab
Conclusions and Directions

• Twitter represents an important platform for e-cigarette/vaping product promotion

• The growth in organic messages followed multi-platform advertising push—advertising works

• Important to capture exposure to formal and informal product promotion across platforms

• Next steps: can we link social media to actual behavior?
Segue: Other Campaigns
HOW DO MESSAGES TRAVEL ACROSS PLATFORMS?
E-cig Tweets and TV Ad Exposure

#HMCTalks
Relationship between TV Exposure (Ratings) and Tweets

CDC Tips 2013

Spearman Rank Correlation 0.83 (p<0.0001)
Pearson Correlation Coefficient 0.74 (p<0.0001)
• Effective to use high-profile events to center marketing strategies
Relationship Between CDC Ads and YouTube Searches

#HMCTalks @hmcollab
Relationship: CDC Ads and Facebook Page Views

#HMCTalks @hmcollab
THINGS WE NEED TO BE THINKING ABOUT
Social Media and Youth

% of All Teens Who Use...

- 71% Facebook
- 52% Instagram
- 41% Snapchat
- 33% Twitter
- 33% Google+
- 24% Vine
- 14% 
- 11% Different Social Media Sites

96% of Youth (13-24) Watching an average of 11.3 hours/week

Source: Pew Center’s Teen’s Relationships Survey
What Do Kids Find Appealing...

This?

Do something today that your future self will thank you for.

Who supports you in your quit journey?

DON'T WISH FOR IT WORK FOR IT

Take the moment make it perfect

Be in love with your life. every moment of it.

Note to self: Every time you were convinced you couldn't go on you did.

If you're still trying you have not failed.

smokefreeus
SmokefreeUs Together we can quit smoking! Join us and share your tips and support to be free of tobacco. http://www.smokefree.gov

528 posts 1,811 followers 58 following
...or This?

Instagram

@hmcollab

#HMCTalks
Social Media Adoption is Widespread

- A low cost health communication strategy
- Missed opportunities for two-way communication
- State tobacco control programs could benefit from more training and guidelines on how to use effectively

<table>
<thead>
<tr>
<th>Platform</th>
<th>2011</th>
<th>2014</th>
<th>Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>28</td>
<td>43</td>
<td>53.60%</td>
</tr>
<tr>
<td>Twitter</td>
<td>23</td>
<td>39</td>
<td>69.60%</td>
</tr>
<tr>
<td>Twitter Followers</td>
<td>2680</td>
<td>9626</td>
<td>259.20%</td>
</tr>
<tr>
<td>Twitter Following</td>
<td>2319</td>
<td>6572</td>
<td>183.40%</td>
</tr>
<tr>
<td>RSS Feeds</td>
<td>7</td>
<td>12</td>
<td>71.40%</td>
</tr>
<tr>
<td>YouTube</td>
<td>14</td>
<td>33</td>
<td>135.70%</td>
</tr>
</tbody>
</table>

@hmcollab
Do Consumers Find State-Sponsored Support?

- Every state has a tobacco control website, but content is static
- SEO is critically important

### Types of sites by tobacco-related search term

<table>
<thead>
<tr>
<th>Tobacco-related Search Term</th>
<th>Public</th>
<th>Private</th>
<th>Public Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quit Smoking</td>
<td>47.90%</td>
<td>29.40%</td>
<td>22.70%</td>
</tr>
<tr>
<td>Tobacco Control</td>
<td>27.40%</td>
<td>1.90%</td>
<td>70.70%</td>
</tr>
<tr>
<td>Secondhand Smoke</td>
<td>34.80%</td>
<td>12.60%</td>
<td>52.60%</td>
</tr>
</tbody>
</table>

#HMCTalks
Media Relationship Summary

- TV matters for tobacco control but Social Media can amplify and extend reach

- Social media host a TON of tobacco promotion and social norms messages...we need to measure this to understand the effects on outcomes.
  - Re outcomes: stay tuned for links with our survey research

- We need to expand definitions of community to include both online communities and local digital communities

- Understanding patterns of diffusion and information propagation can inform campaign strategy and reveal opportunities to leverage investments in traditional media
Acknowledgments/Disclosure

• This research was funded by the National Cancer Institute (NCI grant 1U01CA154254).
• This presentation is the work entirely of the authors and does not necessarily represent the views of the NCI.
• For more information about this research, visit http://healthmediacollaboratory.org/.

#HMCTalks
Stay in Touch

/HealthMediaCollaboratory

@HMCollab
@SherryEmery

bit.ly/bloghmc