SOURCES AND SENTIMENTS OF E-CIGARETTES
AMONG TWITTER USERS

A Thesis
Presented to the
Faculty of
San Diego State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Public Health
with a Concentration in
Health Promotion and Behavioral Science

by
Sharon Nichole Hughes
Summer 2016
SAN DIEGO STATE UNIVERSITY

The Undersigned Faculty Committee Approves the

Thesis of Sharon Nichole Hughes:

Sources and Sentiments of E-Cigarettes Among Twitter Users

[Signature]

Eric R. Buhi, Chair
Graduate School of Public Health

[Signature]

Heather L. Corliss
Graduate School of Public Health

[Signature]

Lourdes S. Martinez
School of Communication

May 6, 2016
Approval Date
ABSTRACT OF THE THESIS

Sources and Sentiments of E-Cigarettes Among Twitter Users
by
Sharon Nichole Hughes
Master of Public Health with a Concentration in Health Promotion and Behavioral Science
San Diego State University, 2016

E-cigarettes have been sited as dangerous to the public’s health because of the carcinogens found in e-liquids, and because of their potential to serve as a gateway drug to combustible cigarettes. The use of e-cigarettes in the US has increased dramatically among youth and adults since 2010. Much of this increase has been linked to the mass marketing that e-cigarette companies have employed through advertisements on television, print, and social media. Currently, the conversations about e-cigarettes on social media are not well understood.

The current study seeks to examine conversations about e-cigarettes on a popular social networking site, Twitter. Tweets related to e-cigarettes were collected using the Social Media Analytic and Research Testbed dashboard. Tweets were collected between October 2015 and February 2016, and a random sample of 1,000 tweets were selected and manually coded for source characteristics and message characteristics. The source characteristics that were measured included: (a) Individuals, (b) Organizations, and (c) Twitter-Verified Celebrities. Within the Individuals and Organizations categories, various details were coded such as whether the user identified as a child, student, parent, potential robot account, business entity, news/media source, or health information provider.

The results of the study showed that a majority of users in the sample identified as Individuals (89.6%) and the remainder identified as Organizations (10.4%). The highest percent of individual accounts discussing e-cigarettes were those coded as Potential Robot Accounts (71.5%) and Children (9.9%). Among Organizations, the highest percent of accounts were coded as Business Entities (75.9%) and News/Media Sources (21.2%). Overall, a majority of profiles posted positive sentiments about e-cigarettes (68.1%). A chi-square test revealed that profiles coded as Organization, are significantly more likely to tweet positive sentiments, compared to Individual profiles: X² (1, N=680) = 639.09. Additional sentiments that were most prevalent included mentions of Stigma (11.7%) and Harmfulness (6.7%). The highest percent of information codes that were discussed include Advertisements (7.2%) and E-Liquid Flavors (7.0%).

Sentiments toward e-cigarettes continue to be very positive. More research should be conducted to investigate the high percentage of potential robot accounts and how these accounts are affecting the conversations happening through social media.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT ............................................................................................................................. iv</td>
</tr>
<tr>
<td>LIST OF TABLES .................................................................................................................. vii</td>
</tr>
<tr>
<td>CHAPTER</td>
</tr>
<tr>
<td>1 INTRODUCTION .........................................................................................................1</td>
</tr>
<tr>
<td>Background ..............................................................................................................1</td>
</tr>
<tr>
<td>The Problem with E-Cigarettes ................................................................................2</td>
</tr>
<tr>
<td>Current E-cigarette Research ...................................................................................4</td>
</tr>
<tr>
<td>Purpose of the Current Study ...................................................................................5</td>
</tr>
<tr>
<td>Strengths and Limitations of the Current Study ......................................................6</td>
</tr>
<tr>
<td>2 LITERATURE REVIEW ..............................................................................................7</td>
</tr>
<tr>
<td>History of E-cigarette Use .......................................................................................7</td>
</tr>
<tr>
<td>E-cigarette Use Today ..............................................................................................8</td>
</tr>
<tr>
<td>E-Cigarettes vs. Combustible Cigarettes ...............................................................10</td>
</tr>
<tr>
<td>Concerns Regarding E-Cigarettes ...........................................................................11</td>
</tr>
<tr>
<td>Current E-cigarette Marketing ..............................................................................14</td>
</tr>
<tr>
<td>E-cigarette Prevention Campaigns ..........................................................................15</td>
</tr>
<tr>
<td>Attitudes and Beliefs About E-cigarettes ................................................................16</td>
</tr>
<tr>
<td>What is Missing From Current Research? .............................................................20</td>
</tr>
<tr>
<td>Significance of Future Research ............................................................................21</td>
</tr>
<tr>
<td>3 METHODS ..................................................................................................................23</td>
</tr>
<tr>
<td>Sample ....................................................................................................................23</td>
</tr>
<tr>
<td>Procedure ...............................................................................................................24</td>
</tr>
<tr>
<td>Instrumentation ......................................................................................................25</td>
</tr>
<tr>
<td>Data Analysis .........................................................................................................27</td>
</tr>
<tr>
<td>4 RESULTS ....................................................................................................................32</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. The E-Cigarette Codebook With Definitions of Each Coding Theme in the Sample...

Table 2. Descriptive Statistics for Twitter Accounts Coded as Individual. (N=896)...........

Table 3. Descriptive Statistics for Twitter Accounts Coded as Organization. (N=104)........

Table 4. Percentage of Tweets That Were Coded Positive, Neutral, Negative, Ambiguous, or Other Tone in the Most Prevalent Source Categories.........................

Table 5. Percentage of Tweets That Discussed Commonly Held Sentiments About E-Cigarettes. (N=1,000) ..............................................................

Table 6. Percentage of Tweets Discussing Specific Subjects About E-Cigarettes..............

Table 7. Descriptive Statistics for Twitter Accounts Coded as Individual, Among Cases That are Not Coded as Potential Robot Accounts. (N=255) ..............

Table 8. Percentage of Tweets That Were Coded Positive, Neutral, Negative, Ambiguous, or Other Tone Among Cases That Were Not Coded Potential Robot Account. ....................................................
CHAPTER 1

INTRODUCTION

Electronic cigarettes (e-cigarettes) are new devices that deliver nicotine to their user. In this chapter I will present the background of both combustible cigarettes and e-cigarettes, and discuss why e-cigarettes pose a possible threat to public health in the United States. The current research of e-cigarettes will be discussed, and I will propose how the current study will contribute to e-cigarette research.

BACKGROUND

Combustible cigarettes are known to cause respiratory diseases such as lung cancer and Chronic Obstructive Pulmonary Disease (COPD), and tobacco use is the leading cause of preventable death in the US (U.S. Department of Health and Human Services [HHS], 2015). Smoking cessation, prevention interventions, and advertisement campaigns have been implemented to decrease the public’s use of combustible cigarettes in the US for over 40 years (Centers for Disease Control and Prevention [CDC], 2006). Interventions that are targeted toward changing tobacco policy and environments have since proven to be most effective in reducing smoking rates (Report of the Surgeon General, 2012). One effective intervention that changes both environments and policy, is restricting tobacco advertising. Prevalence of tobacco advertisement in a person’s environment directly affects the likelihood that they will buy cigarettes (National Cancer Institute, 2008). Because of this link, tobacco advertising has been banned or restricted in various locations and channels throughout the US (Maloney & Cappella, 2015).

In 2007, e-cigarettes were first introduced in the United States (US) market, and have been increasing in awareness and use ever since (McMillen, Gottlieb, Whitmore-Shaefer, Winickoff, & Klein, 2014). E-cigarettes are also referred to as e-cigs, vape, vaping, e-pens, and e-vape. The e-cigarette is a nicotine delivery device that has been regarded as a safer
alternative to traditional, combustible cigarettes because they do not expose their users to the same amount or degree of toxins as traditional cigarettes (Cahn & Siegel, 2011). Although the device mimics a cigarette, and contains many of the same carcinogens, e-cigarettes are not considered tobacco products, and thus have not been regulated by the same marketing policies (Maloney & Cappella, 2015). In June 2016, however, this will change as the Food and Drug Administration (FDA) finalized a rule to regulate the manufacturing, advertising, promotion, and sale of e-cigarettes (FDA, 2016).

E-cigarettes present public health professionals with a difficult situation because their stance on this product forces them to choose between tobacco prevention and harm-reduction (DeLuca, 2016). DeLuca (2016) declared that the e-cigarette debate has split tobacco researchers into two groups (those for and those against e-cigarettes) and thus has made the voice of public health unclear. The argument for harm-reduction states that e-cigarettes may be the most promising form of tobacco reduction because e-cigarettes allow smokers to engage in the rituals of smoking without exposing the user to combustible products (Polosa, Rodu, Caponnetto, Maglia, & Raciti, 2013). The American Public Health Association (APHA), however, has taken the alternative stance in support of preventing the use of e-cigarettes by advocating a tax on e-cigarettes similar to tobacco products and tasking the FDA to regulate their advertisement, production, and sale (APHA, 2016). APHA (2016) argued that the evidence for e-cigarettes as a smoking cessation tool is not conclusive, and that the widespread use of e-cigarettes will institute more harm than reduction. A major reason for this argument is that e-cigarette use has been rapidly increasing among youth and there is evidence that suggesting that e-cigarette use may lead to minors trying other tobacco products, such as combustible cigarettes (APHA, 2016). Although the voice of the harm-reduction argument was greatly promoted by e-cigarette companies and some health professionals, the suggestions made by APHA will be implemented beginning in June 2016 (FDA, 2016).

**The Problem with E-Cigarettes**

E-cigarette use has been gaining popularity among youth in the US since 2013. Among high school students, current use of e-cigarettes nearly tripled between 2010 and 2014 (4.5% versus 13.4%). Current use among middle school students more than tripled
between 2010 and 2014 (1.1% vs. 3.9%) (CDC, 2015). Experts have largely attributed increased awareness and use of e-cigarettes among youth to the marketing efforts of e-cigarettes, such as advertising decorative e-pens and offering creatively flavored juices that are smoked when vaping (Rath, Villanti, Abrams, & Vallone, 2012). Researchers and professionals see increased e-cigarette use as a potential danger for three reasons: (a) the carcinogens in e-cigarettes may pose a health threat to e-cigarette smokers, (b) they may serve as a *gateway drug* to traditional cigarette use, and (c) the use of e-cigarettes may re-normalize smoking behavior (Fairchild, Bayer, & Colgrove, 2014; Leventhal et al., 2015; Morris et al., 2015; Primack, Soneji, Stoolmiller, Fine, & Sargent, 2015; Stanwick, 2015).

The first concern, regarding the carcinogens found in e-cigarettes, has not been studied extensively. Currently, little research has been conducted that reports the short-term effects of using e-cigarettes, and there are no studies that report the long-term effects of e-cigarette use. The carcinogens of concern in e-cigarettes include nicotine, carbonyls, and particulate matter (Kosmider et al., 2014; Morris et al., 2015; Turner et al., 2015). Short-term effects of respiratory exposure to these carcinogens include irritation of the mouth, throat, and lungs (Bekki et al., 2014). Regardless of the lack of research, e-cigarettes are identified as a potential threat to health because of the potentially harmful effects of nicotine and particulate that is exposed to the e-cigarette smoker (Morris et al., 2015).

The second concern, regarding the possibility of e-cigarettes serving as a gateway drug to traditional cigarettes, has just begun to be studied. In fall of 2015, two studies were published that examined this theory. These studies were self-report surveys that included teens and young adults from across the US. The studies sought information about the participants’ use of combustible cigarettes before and after they began using e-cigarettes. Both studies found that adolescents and young adults who used e-cigarettes (and had never smoked cigarettes) were significantly more likely to begin smoking regular cigarettes one year later (Leventhal et al., 2015; Primack et al., 2015). Because there is evidence that e-cigarettes may serve as a gateway drug to combustible cigarettes, it is crucial to continue studying e-cigarettes further.

The third concern is that the increased use of e-cigarettes will re-normalize smoking behavior. E-cigarettes mimic combustible cigarettes in their form and function, they look similar to cigarettes, and are smoked the same way that cigarettes are smoked (Stanwick,
Studies that have contributed to this idea, show that most e-cigarette users believe that e-cigarettes do not have the same negative stigmas that combustible cigarettes have (Coleman et al., 2015), and e-cigarettes can be used in most public areas (Dawkins, Turner, Roberts, & Soar 2013; Pepper, Ribisl, Emery, & Brewer, 2014).

**CURRENT E-CIGARETTE RESEARCH**

An innovative way that researchers have begun to study e-cigarettes is through public Twitter accounts. The aspects of e-cigarettes that have been studied through Twitter are: (a) source characteristics (studies reported what type of accounts were discussing e-cigarettes), (b) sentiments (studies reported if Twitter users were discussing e-cigarettes with positive, negative, or neutral sentiments), and (c) the information discussed in each sample (studies reported if the sample contained details about advertisements, flavors, or smoking cessation) (Cole-Lewis et al., 2015). Thus far, limited studies have been conducted in which Twitter content (*tweets*) are analyzed for the public’s sentiments and beliefs about e-cigarettes (Cole-Lewis et al., 2015; Godea, Caragea, Bulgarov, & Ramisetty-Mikler, 2015; Huang, Kornfield, Sczypka, & Emery, 2014; Myslin, Zhu, Chapman, & Conway, 2013). Each of these studies found that the majority of Twitter users share information or advertisements concerning e-cigarettes. Further, small portions of the tweets discuss opinions about e-cigarettes, which are mostly positive (Cole-Lewis et al., 2015).

When data were collected for the Twitter data studies, widespread awareness campaigns of the potential dangers of e-cigarette use had not been launched. In March 2015, this changed when the Centers for Disease Control and Prevention (CDC) began releasing their first educational ads about the dangers of e-cigarettes (HHS, 2015). Shortly after this, the California Department of Public Health (CDPH) launched a campaign called Still Blowing Smoke that warns the public about misleading information circulating about the safety of e-cigarettes (CDPH, 2015). Since then, the San Francisco Department of Health has also implemented the #CurbIt campaign to warn the public about the potentially harmful and addictive properties of vaping (San Francisco Tobacco-Free Project, 2015), and the Chicago Department of Public Health began the #VapingTruth campaign, which aims to reduce youth smoking through education about the toxic substances found in e-liquids (Healthy Chicago, 2015).
Studies have not been conducted to reflect public sentiments and beliefs since these awareness campaigns have launched. In order to gain current knowledge on the sentiments and opinions of people in the US about e-cigarettes, data should be collected after the campaigns’ launch. Because there is no conclusive research that highlights the effects of e-cigarette use, another need in future research are studies that examine the short-term and long-term effects of e-cigarette use. The studies that used Twitter data agreed that analyzing tweets for information about respiratory health effects among e-cigarette users is the next step in providing meaningful information to the literature (Godea et al., 2015; Myslin et al., 2013).

**PURPOSE OF THE CURRENT STUDY**

The primary purpose of the proposed research study is to build on the existing knowledge about sentiments, source information, and opinions about e-cigarette use of Twitter users. After national, statewide, and local campaigns were implemented to prevent e-cigarette use, the proposed study will assess the current sentiments about electronic cigarettes among Twitter users. The study will also report common themes about Twitter users’ experience with e-cigarettes, and information about the source of the tweets. Lastly, the study will report the percent of Twitter users that are reporting respiratory health effects associated with e-cigarette use.

In the current study, I will use Social Media Analytic and Research Testbed (SMART) technology to gather tweets from Twitter users who are discussing topics about e-cigarette use in the US (Tsou, Jung, et al., 2015). The SMART dashboard is a software that allows its user to input topic-specific keywords into the program, and over a specified period of time the software will yield relevant tweets from an application program interface (API) from publically available tweets (Tsou, Peddecord, Johnson, & Jung et al., 2015). I will code a randomly selected sample of 1000 of the collected tweets in order to test for themes regarding source characteristics, sentiments, and information. Analyzing tweets for these themes will give researchers a preliminary view of who is discussing e-cigarettes and their views about e-cigarettes. The proposed research will also provide public health and communication professionals with useful information that will aid them in designing awareness campaigns and interventions around the prevention or cessation of e-cigarette use.
STRENGTHS AND LIMITATIONS OF THE CURRENT STUDY

There are many strengths in the proposed research. A major strength of this research is that the researcher is able to collect a large amount of data in real time. The data used for this study is current (it was collected from October 2015 to February 2016) and the sampling frame includes all geographic locations in the US.

This proposed study is not devoid of limitations. A limitation to this research is that demographic and geographic information cannot be validated. All demographic data and content will be collected from Twitter profiles that may or may not contain true information about the sender’s geographic location, age and other characteristics. More detailed demographic information such as ethnicity and sex cannot reliably be collected from Twitter profiles because users often choose not to display this information. It cannot be verified that the SMART dashboard will collect a representative sample of people in the US or on Twitter, since the technology only collects a specific percentage of available tweets. Lastly, additional background information cannot be gathered about each Twitter user. Useful information that cannot be collected in this study includes past or current use of traditional cigarettes.
CHAPTER 2

LITERATURE REVIEW

In this section I will highlight the specifics of e-cigarette use in the US, how the contents of e-cigarettes compare to the dangers of combustible cigarettes, and how they pose a threat to public health. I will then discuss the current efforts of e-cigarettes companies to target youth in their marketing, how this marketing has shaped the attitudes and beliefs about e-cigarettes, and how public health departments are attempting to mitigate these effects. Lastly, I will discuss what is missing from the current e-cigarette research and the importance of future research in this area.

HISTORY OF E-CIGARETTE USE

E-cigarettes were first introduced to the US market between 2006 and 2007 (Consumer Advocates for Smoke-free Alternatives Association, 2012). The e-cigarette has since been regarded as a safer alternative to traditional, combustible cigarettes because they deliver nicotine to user while exposing them to fewer toxins compared to traditional cigarettes (Cahn & Siegel, 2011; Goniewicz et al., 2014). E-cigarettes have been widely regarded and marketed by their supporters as a tool to assist traditional cigarette smokers with smoking cessation (noted by 64 percent of e-cigarette websites), and a product that does not expose others to second-hand smoke (noted by 22% of e-cigarette websites) (Consumer Advocates for Smoke-free Alternatives Association, 2012; Grana & Ling, 2014).

E-cigarettes are smoked through e-pens or vaporizers (also known as vapes), which vary greatly in design and performance (Brown & Cheng, 2013). All e-cigarettes use a power source (such as a battery) and a heating device to deliver the e-liquid as an aerosol mist to the user. These e-liquids contain nicotine, flavorings, and other carcinogens (Glasser et al., 2015). The e-cigarette delivers the e-liquid vapor to the user’s respiratory system when the
user inhales through the cartridge, and the battery is activated (Caponnetto, Polosa, Russo, Leotta, & Campagna, 2011).

Today, health professionals, government agencies, and the general public debate the e-cigarette’s potential to help traditional cigarette smokers quit or prevent relapse, and to serve as a gateway drug to people who have never smoked combustible cigarettes (Bell & Keane, 2014; Leventhal et al., 2015; Morris et al., 2015; Pentz et al., 2014; Primack et al., 2015). E-cigarettes are not regulated or approved by the Food and Drug Administration (FDA) as a smoking cessation tool, and are prohibited to be advertised as such. In this review, I will first provide an overview of research on e-cigarette prevalence and awareness, as well as the health concerns and potential effects of e-cigarettes. The role of marketing of e-cigarettes will then be explained, followed by a description of the health campaigns that both promote and warn users about the effects of e-cigarette use. Lastly, current research regarding the US population’s sentiments and knowledge about e-cigarettes will be reviewed, as well as next steps for research to improve understanding of gaps in knowledge about e-cigarettes. In this review, the act of smoking e-cigarettes will be referred to as vaping and e-cigarette users will be referred to as vapers.

**E-Cigarette Use Today**

E-cigarette use has been rapidly increasing in awareness and use among both youth and adults in the US since nationally representative data began being collected on e-cigarette use in 2013 (CDC, 2015). The CDC reported that e-cigarette use from 2010-2014, among adolescents who had used e-cigarettes on at least one of the last 30 days, has increased from 4.5% to 13.4% in high schools and has increased from 1.1% to 3.9% in middle schools. As of 2014, current use of e-cigarettes has surpassed current use of all other tobacco products for the first time since data collection began in 2011 (CDC, 2015). E-cigarettes have also been gaining popularity in adult populations (CDC, 2014). The CDC reported that in 2013, the percentage of adults who had ever used e-cigarettes had more than doubled, from 3.3% to 8.5%. Current use of e-cigarettes has more than doubled from 1% to 2.5% from 2010 to 2013. In the same time frame, awareness of e-cigarettes also doubled from 40% to 79%; this is attributed mostly to the heavy marketing surrounding e-cigarettes (CDC, 2014; King, Patel, Nguyen, & Dube, 2014). Another study found even higher
instances of prevalence in adults who have ever used e-cigarettes and are currently using e-cigarettes (McMillen et al., 2014). McMillen et al. (2014) conducted a study of over 3,000 participants and examined the demographic trends of e-cigarette use among US adults from 2010 to 2013 by using a cross-section design and a nationally representative sample in each of the four years that the study was conducted. The study found that current use among adults increased from 0.3% (in 2010) to 6.8% (in 2013), and ever use of e-cigarettes among adults increased from 1.8% (in 2010) to 13.0% (in 2013). Among young adults, aged 18-24 years (14.2%), was higher than adults aged 25-44 years (8.6%), 45-64 years (5.5%), and 65 years and older (1.2%). The survey also found that in 2013, daily cigarette smokers and nondaily cigarette smokers were most likely to report that they currently use e-cigarettes, but 32% of people who reported currently using e-cigarettes had never smoked traditional cigarettes (McMillen et al., 2014).

There are also gender and ethnic differences of e-cigarette use (Littlefield, Gottlieb, Cohen, & Trotter, 2015; Shoenborn & Gindi, 2014). In 2014, the National Health Interview Survey (NHIS) found that men were more likely than women to have ever tried e-cigarettes (14.2% vs. 11.2%) and to report current use of e-cigarettes (4.1% vs. 3.4%) (Shoenborn & Gindi, 2014). In regard to ethnicity, adults in the US who are most likely to report current use of e-cigarettes are American Indian or Alaskan Native (10.7%), followed by Non-Hispanic White (4.6%), and Hispanic (2.1%). The ethnicity that reported the lowest current use of e-cigarettes was Non-Hispanic Black (1.8%) and Non-Hispanic Asian (1.5%) (Shoenborn & Gindi, 2014).

Heightened awareness and use of e-cigarettes among adolescents is largely attributed to the marketing efforts of e-cigarette manufacturers. Manufacturers are using tactics such as advertising decorative e-pens and offering creatively flavored juices that are smoked when vaping, to make e-cigarettes more appealing to youth (Rath et al., 2012). Cigarette companies consider package branding as a critical component of their overall marketing strategy (Wakefield, Worley, Horan, & Cummings, 2002). The idea of branding cigarettes to appeal to young people has been tested and has been proven effective (Hammond, Doxey, Daniel, & Bansal-Travers, 2011). Specifically, Hammond et al. (2011) conducted a study in which groups young women (18-19 years, smokers and non-smokers) were offered either a pack of plain cigarettes or a branded female pack of cigarettes. The female-branded packages
included one or more of the following features: (a) thin cursive writing for prominent titles, (b) slimmer packaging than the traditional sized package, (c) greater amounts of white space with thin smoke-like designs, and (d) use of non-traditional smoking colors (pink, purple, and blue). The results showed that the female-branded packs were more likely to be perceived as less harmful and were three times more likely to be requested at the end of the study by both smokers and non-smokers, than were plain packs (Hammond et al., 2011). Because branding cigarettes to appeal to a specific gender is effective in making the target audience desire the cigarettes significantly more, despite their initial interest in smoking, there may be evidence that non-smokers are more likely to try a cigarette or e-cigarette that is branded to appeal to them.

**E-Cigarettes vs. Combustible Cigarettes**

The link between smoking combustible cigarettes and its cause of multiple cancers and health effects, such as chronic obstructive pulmonary disease (COPD), is well supported by current research (Levitz, Bradley, & Golden, 2004). Many carcinogens that are found in combustible cigarettes are also found in e-cigarettes, which is why many health professionals are skeptical as to the health benefits that e-cigarettes can provide (Morris et al., 2015). Although there are no longitudinal studies available to lend insight on the long-term health effects of e-cigarettes, there is available evidence that has tested their safety as compared to combustible cigarettes (Farsalinos, Kristler, Gillman, & Voudris, 2015; Goniewicz et al., 2014; Misra, Leverette, Cooper, Bennett, & Brown, 2014).

In a study conducted by Goniewicz et al. (2014) and colleagues, researchers analyzed the contents of e-cigarette vapor from 11 popular e-cigarette brands in order to compare their contents to traditional cigarettes. Fifteen carbonyl compounds were tested in the vapor (including formaldehyde, acetaldehyde, valeric aldehyde, and hexanal) and twelve metals (including cadmium, nickel, and lead). The results showed that four of the carbonyls were found in e-cigarette vapor and three metals were found. Each of the substances found in the e-cigarette vapor is known to be toxic to humans (World Health Organization [WHO], 2015). Although toxic substances were found in the e-cigarette vapor, as compared to cigarette smoke, it was much less. Specifically, formaldehyde was found at a concentration of nine times less than cigarette smoke, acetaldehyde was found at a concentration of 450 times less.
than cigarette smoke, and acrolein was measured to be at a concentration of 15 times less than what is found in cigarette smoke (Goniewicz et al., 2014).

To further determine the contents of e-cigarette vapor from different e-liquid brands, a study conducted by Misra et al. (2014) and colleagues, collected 12 different brands of e-cigarette vapor, and then tested them for prevalence of toxic substances. The study concluded that the concentration of toxic substances found in e-cigarettes was very limited compared to combustible cigarettes. The study also found that the nicotine concentration in e-cigarettes was 10 to 20 times higher than in traditional combustible cigarettes (Misra et al., 2014).

Because of the high nicotine concentration in many e-liquids, there is high risk of addiction to nicotine by using e-cigarettes. Also, although e-cigarettes contain significantly lower concentrations of toxins than combustible cigarettes, it is still unknown as to whether the levels of toxins that do exist are of significant concern.

**CONCERNS REGARDING E-CIGARETTES**

Researchers and professionals see increased e-cigarette use as a potential danger for two major reasons: (a) the carcinogens in e-cigarettes may pose a health threat to e-cigarette smokers, and (b) they may serve as a gateway drug to traditional cigarette use for non-smokers (Leventhal et al., 2015; Morris et al., 2015; Primack et al., 2015). Further, public health professionals fear that e-cigarettes will re-normalize smoking behavior and the general public will no longer see smoking as a negative activity (Fairchild et al., 2014; Stanwick, 2015).

The first concern, regarding the carcinogens found in e-cigarettes, has not been studied extensively. Currently, little research has been conducted that reports the short-term and long-term effects of using e-cigarettes alone. The carcinogens in e-cigarettes that are of primary concern are nicotine, carbonyls, and particulate matter (PM). Each of these contents is also found in traditional cigarettes (Morris et al., 2015).

Nicotine is considered safer than tobacco, but nicotine itself causes many health hazards. Studies have shown that some of these health risks include increased risk of cardiovascular, respiratory, and gastrointestinal disorders, reproductive issues, and DNA mutation, which leads to cancer (Mishra et al., 2015). The health effects of short-term use of
nicotine are not of great concern, but prolonged use of nicotine is of moderate concern (Morris et al., 2015).

The second carcinogens that are of concern in e-cigarettes are carbonyls. Carbonyls, such as formaldehyde, acetaldehyde, and acrolein, have been found in varying degrees of e-cigarette vapors and are known to cause dry mouth and dry throat of vapers (Bekki et al., 2014). The solvent used by the e-liquid and the battery output voltage determines the level of carbonyls in e-cigarette vapor. One study found that high-voltage devices yielded carbonyl levels similar to that of tobacco smoke (Kosmider et al., 2014). The e-cigarette user’s choice to smoke with a high-voltage device, and their ability to use a low-voltage device, are important factors to consider when determining the health effects of e-cigarettes.

Lastly, PM is a well-known carcinogen that, when exposed, increases risk of lung cancer. Particulate matter is found in the ambient air, and increases risk of lung cancer mortality independent of smoking cigarettes (Turner et al., 2015). These carcinogens are found in even greater concentrations in traditional and electronic cigarettes, than in ambient air (Li, Li, Bai, & Song, 2015). Because there are potentially dangerous amounts of PM being exposed to e-cigarette users, it is reasonable to consider PM an agent of concern in e-cigarettes.

The second concern, regarding the possibility of e-cigarettes serving as a gateway drug to traditional cigarettes, has just begun to be studied but there are three recent studies that give significant insight on the subject. It is especially important to examine the use of e-cigarettes among youth and their potential to serve as a gateway to traditional cigarettes because in the 2012 Surgeon General’s Report it was found that 89% of current smokers first tried smoking in their teens, and 75% of teen smokers continued smoking as adults (HHS, 2014). Evidence of the claim that e-cigarettes could serve as a gateway to traditional cigarettes, was documented in one study that analyzed data collected from young adults (age 18-29 years) who participated in the National Adult Tobacco Survey. The study examined the responses of participants who reported that they had never developed a cigarette smoking behavior, and how that related to their self-reported intention not to smoke in the next year. Results from the study found that ever use of e-cigarettes and other tobacco products, was associated with less intention to not smoke in the next year (or more openness to smoking traditional cigarettes) (Coleman et al., 2014).
To lend further evidence to the gateway drug debate, in fall of 2015, two studies were published that examined the idea of e-cigarettes being used as a gateway to cigarette use. Both studies found that adolescents and young adults who used e-cigarettes (and who had never smoked cigarettes) were significantly more likely to develop a habit of smoking combustible cigarettes one year later (Leventhal et al., 2015; Primack et al., 2015). Primack et al. (2015) conducted a longitudinal study with 694 teens and young adults across the US who were not, and had never been, traditional cigarette smokers at baseline. Data were collected at baseline and one year later, and participants self-reported their e-cigarette and cigarette use on a validated survey. Researchers analyzed whether the e-cigarette users were more likely to develop a traditional cigarette smoking habit than the non-e-cigarette users. Results showed that nearly 69% of e-cigarette users developed a smoking habit one year later, while only 19% of participants who did not use e-cigarettes developed a smoking habit.

Leventhal et al. (2015) also conducted a longitudinal study in which over 2500 youth (14 years old) in Los Angeles who had never used traditional cigarettes completed validated surveys at baseline, six, and twelve months regarding their e-cigarette and cigarette use. The results found that if a youth had reported that they have ever used e-cigarettes, they were significantly more likely to have later used cigarettes in the last six months, than those participants who had never used e-cigarettes. After 12-month follow-up, over 25% of youth who reported using e-cigarettes, also reported smoking combustible cigarettes. Conversely, only 9% of youth that reported ever use of e-cigarettes had used combustible cigarettes. Leventhal et al. (2015) and Primack et al. (2015) both controlled for a number of risk factors of smoking such as, having friends or parents that smoke.

Lastly, Willis et al. (2016) conducted a similar longitudinal study that specifically studied over 2000 ninth and tenth graders in Hawaii. This study included a survey that assessed e-cigarette use at baseline, and combustible cigarette use one year later. The study found that adolescents, who used e-cigarettes at baseline, were significantly more likely to report smoking combustible cigarettes within the last 30 days, than adolescents who did not report using e-cigarettes at baseline (Willis et al., 2016). These studies serve as crucial evidence for the importance of why the health effects of e-cigarettes should continue to be studied.
In addition to the above stated major concerns that make e-cigarettes potentially harmful, the technology and make up of the e-cigarette itself poses harm to e-cigarette users (Chen, 2013). Chen (2013) conducted an adverse effects surveillance test of e-cigarettes and submitted it to the FDA. The surveillance test reported that there were many cases in which the batteries in e-cigarettes blew up in the faces or in the pockets of recreational e-cigarette users. These accidents caused second-degree burns on the faces, hands, and legs, of e-cigarette users.

**CURRENT E-CIGARETTE MARKETING**

When discussing the potential impact of e-cigarettes, it is important to consider the environment in which they exist. E-cigarette marketing has been very prevalent in the US because their advertisement has not been regulated by the FDA (Duke et al., 2014). Prior to June 2016, because e-cigarettes do not contain tobacco, the only way that the FDA could have potentially regulated e-cigarette advertisement is if they were being marketed as a smoking cessation tool (Duke et al., 2014). E-cigarettes are not approved by the FDA as a smoking cessation tool though, so legally the product could not be advertised as such, and it’s marketing could not be regulated as a tobacco product. During the course of this study, e-cigarette companies were free to market their products on television and throughout stores in the US. E-cigarette advertisements were also not restricted among youth populations (Maloney & Cappella, 2015). Beginning in June 2016, e-cigarettes will be regulated completely by the FDA, and widespread marketing will be prohibited (FDA, 2016).

As was mentioned earlier, the rise in awareness and use of e-cigarettes is attributed greatly to the unregulated advertisement of these products. The 2014 National Youth Tobacco Survey reported that youth exposure to e-cigarette advertising is most prevalent in retail stores, but exposure is wide-reaching; 14.4 million youth were exposed in retail stores, 10.5 million youth were exposed on the internet, 9.6 million youth were exposed through television or movies, and 8 million youth were exposed through newspapers or magazines (CDC, 2014). One study showed that in television marketing alone, youth exposure to e-cigarette advertisements increased by 256% from 2011 to 2013, and young adult exposure to e-cigarette advertisement increased by 321% during the same time frame (Duke et al., 2014). Until March 2015, the vast majority of e-cigarette advertisements relayed a message of
potential health benefits and safety that they offer. There were no national, state, or local prevention campaigns that disparaged e-cigarette use or warned users of the potential harm that they could cause.

**E-CIGARETTE PREVENTION CAMPAIGNS**

Since spring of 2015, four large-scale anti-vaping marketing campaigns have been launched in the US: (a) **TIPS**, (b) **Still Blowing Smoke**, (c) **#CurbIt**, and (d) **#VapingTruth**. In March 2015, the Centers for Disease Control and Prevention (CDC) began releasing their first educational advertisements about the dangers of e-cigarette use as a part of their **TIPS** campaign (HHS, 2015), and distributed the materials on a national scale. The messages included statements such as, “I started using e-cigarettes but kept smoking. Right up until my lung collapsed.” The CDC also dedicated a page on their website to information and warnings about e-cigarettes, and has released a series of educational webinars called **Vitalsigns** regarding youth exposure and advertisement of e-cigarettes.

Shortly after the CDC released their campaign materials, the California Department of Public Health (CDPH) launched a campaign called Still Blowing Smoke, that warns the public about misleading information that has been circulating about the safety and benefits of e-cigarette use (CDPH, 2015). The Still Blowing Smoke campaign aims to inform the public of the potential dangers of e-cigarettes, how e-cigarettes pose a direct threat to youth, and provide clarity about Big Tobacco’s involvement in the industry.

The San Francisco Department of Public Health then employed their campaign called **#CurbIt**. This campaign aims to dispel myths about e-cigarettes such as “you can smoke e-cigarettes anywhere” (San Francisco Tobacco-Free Project, 2015). The **#CurbIt** campaign is primarily based online, and gives youth a chance to participate in the campaign by using the hashtag **#CurbIt**, on social media posts. The campaign has materials with messages such as, “E-cigarettes are harmful, like cigarettes. Not allowed in public. Vape and smoke only at the curb.” This campaign also blames Big Tobacco for marketing nicotine-containing devices to youth.

Lastly, in December 2015, Chicago launched an anti-vaping campaign called **#VapingTruth**. It was a part of the Health Chicago initiative that is spearheaded by the Chicago Department of Public Health. Although the campaign does have an online presence,
much of the material is in the form of billboards and posters that are distributed throughout Chicago (Healthy Chicago, 2015). The campaign aims to reduce youth smoking by educating about the highly addictive nature of nicotine, toxic chemicals found in e-liquid, and the lack of regulation of e-cigarette devices. Campaign messages include, “Vaping: liquid poison,” and “Vaping: it’s still addiction.”

These campaigns were the first in the US to widely address the potential dangers of e-cigarette use for smokers and non-smokers. Because these campaigns were implemented recently, there is currently no research on attitudes and beliefs of the US public about e-cigarettes since the release of health awareness campaigns.

**ATTITUDES AND BELIEFS ABOUT E-CIGARETTES**

It is important to consider the opinions and sentiments of the populations that are exposed to e-cigarette advertisement, because this may be an indicator of their intention to use e-cigarettes. An innovative way in which researchers have gained information about sentiments concerning e-cigarettes is from social media accounts. In particular, there have been many studies that captured the public’s sentiments, knowledge, and experience about different health subjects by gathering Twitter data (Alvaro et al., 2015; Krauss et al., 2015). These studies included capturing sentiments and experience regarding adverse prescription drug reactions (Alvaro et al., 2015), hookah use (Krauss et al., 2015), obesity (Kent et al., 2015), and more.

Twitter is an online social networking and microblogging site in which registered users post, share, and read short messages (140 characters or less), pictures and videos. These messages are called tweets and can be retweeted (re-posted by another user), favorited (acknowledged favorably by another user), or responded to by another user. As of September 2015, Twitter has over 307 million active users across the world (Statista, 2015). Specifically in the US, 23% of all Internet users use Twitter, and 20% of the whole adult population has a Twitter account. Thirty-two percent of young adult Internet users (ages 18 to 29 years) have a Twitter account, and 29% of Internet users ages 30-49 have a Twitter account (Pew Research Center, 2015). Among those who have Twitter accounts 22.6% are under age 24, and 21.5% of users are 25-34 years (Pew Research Center, 2015). The Pew Research Center (2015) reports that 59% of Twitter account holders use Twitter at least weekly. Because Twitter is
very popular among this population, it is an effective way to reach young demographics is the US.

A key way in which Twitter data is helpful in health research is that tweets can contain rich detail that give researchers insight into various health topics (Jimeno-Yepes, MacKinlay, Han, & Chen, 2015). Recently, researchers have begun to collect tweets in real-time, and use this information for biosurveillance. Biosurveillance is a tactic in which health professionals gather data based on location in order to monitor the spread of information and disease (National Association of County & City Hall Officials [NACCHO], 2016). Researchers have been able to use Twitter to conduct biosurveillance of the Ebola outbreak and Influenza (Odlum & Yoon, 2015; Santillana, Nguyen, Dredze, Paul, & Brownstein, 2015).

Although Twitter is known to be effective in retrieving pertinent health information, relatively few studies have been conducted in which public Twitter accounts are analyzed for sentiments and information about vaping. In addition, no studies have been conducted since the various anti-vaping campaigns were implemented in March 2015 (Cole-Lewis et al., 2015; Godea et al., 2015; Huang et al., 2014; Myslin, Zhu, Chapman, & Conway, 2013). One study that collected tweets from around the US, analyzed Twitter data to report general sentiments and information dissemination about e-cigarettes (Godea et al., 2015). The study collected over 105,000 tweets over the span of two months in 2014 using seven keywords about e-cigarettes. A sentiment and informational analysis was conducted over a random sample of 1,000 tweets. Sentiments were measured by codes labeled Positive, and Negative, and content was classified as Advertisement, Informational, and Other. The study found that 33% of tweets were information sharing, 22% of tweets were advertising e-cigarettes, 11% of tweets contained positive sentiments about e-cigarettes, and 3% of tweets contained negative sentiments about tweets (Godea et al., 2015). One limitation of this study is that the categories, such as Informational, are very broad and do not give the reader much detail about the informational tweet that was coded. For example, the tweet may be about smoking cessation, policy, or safety, but the broad categories do not give the reader those specifics.

Clark et al. (2016) conducted a study in which e-cigarette related tweets were coded as Automated or Organic. The Automated tweets measured in this study were primarily focused on e-cigarette marketing or were promotional in nature. The results of the study
found that 80% of tweets in the sample were automated, which closely translated to an advertisement coded in other studies. The percentage of advertisement tweets that were found in the sample (80%) was much larger than what was found in previous studies (Clark et al., 2016). A major weakness of this study is that it does not report what kind of profiles are tweeting these automated tweets, therefore public health professionals do not know what kind of profiles are disseminating such advertisements.

An additional study (Myslin et al., 2013) that measured content and sentiments of tweets regarding e-cigarettes, used keywords to collect over 7,000 tweets over 15 day intervals from December 2011 to July 2012. This study also measured positive and negative sentiments, as well as the content of the tweet including first-hand experience, informational, news, marketing, cessation, social relationships, and more. The study found that the most common themes were first-hand experience, second-hand experience, and opinion. Overall, sentiments were more positive (46%) than negative (32%) (Myslin et al., 2013). One strength of this study is that tweets were collected over a longer period of time, and the categories that were used provided the reader with more detail about the information shared in the tweet. One limitation to this study is that the opinions that were gathered were not outlined in detail, so the public’s beliefs about e-cigarettes were not presented.

Huang et al. (2014) conducted a study in which e-cigarette-related tweets were collected by using keywords, and then manually categorized based on content. This study also made use of broad categories such as smoking cessation mentions, health and safety mentions, and advertising mentions. Over 73,000 tweets were collected by using a service called the Twitter Firehose. Twitter Firehose is a service that streams all public data that are posted on Twitter, and gives researchers access to all public Twitter statuses that fit their search criteria (Twitter Firehose, 2016). This study found that over 90% of tweets collected were advertising e-cigarettes or e-cigarette materials. The results lend insight into the vast amount of advertising that the public is exposed to in regard to e-cigarettes.

Cole-Lewis et al. (2015) conducted a similar study in which 17,000 tweets were collected using e-cigarette-related keywords, and analyzed them for knowledge, sentiments, and opinions. The goal of the study was to test a machine-learning system that could code tweets appropriately based on their content. Using the machine-learning technology, the tweets were coded into five different categories including relevance, user description, genre
and theme. Within each of these five categories, multiple sub-categories were coded in order to give additional detailed information about the tweet. The results showed that a machine-learning technology could be effectively used to code tweets correctly, which is much faster than humans manually coding tweets (Cole-Lewis et al., 2015). A limitation of the study is that the results focused on the effectiveness of machine-learning system in coding tweets correctly, and did not report the data that was retrieved regarding sentiments, content, and information.

Each of these studies are noteworthy because they were able to analyze public information, sentiments, and source characteristics in real-time. They included thousands of tweets from users across the US, and contain unfiltered public opinion. The studies also effectively obtained information from thousands of Twitter users in a short period of time. E-cigarette research is an appropriate topic to gain Twitter data from because it is a rapidly evolving industry, and it will be helpful to researchers to quickly gather data to understand the public’s changing opinion over time (Cole-Lewis et al., 2015).

Along with the Twitter data studies that measured sentiments about e-cigarettes, a separate study measured sentiments and opinions about e-cigarettes of over 1,400 cigarette smokers through an online survey (Harrel et al., 2014). This study compared smokers’ sentiments about e-cigarettes vs. traditional cigarettes. The researchers found that smokers believed that e-cigarettes had fewer health risks and caused fewer cravings, less addiction, and fewer negative physical feelings than combustible cigarettes (Harrel et al., 2014). These positive sentiments and beliefs about e-cigarettes are consistent with results obtained from Twitter data studies (Godea et al., 2015; Myslin et al., 2013). Other studies that have collected survey information about peoples’ sentiments and opinions about e-cigarettes found that positive attitudes about e-cigarettes included e-cigarette ability to assist in smoking cessation, less exposure to second-hand smoke when using e-cigarettes, ability to use e-cigarettes in most places, and e-cigarettes carrying fewer stigmas than traditional cigarettes (Coleman et al., 2015; Dawkins et al., 2013; Farsalinos, Kristler, Gillman, & Voudris, 2015; Pepper et al., 2014; Zhu et al., 2014). The percent of these sentiments and opinions will be measured in the proposed research study. Without health education or research on the effects of e-cigarettes, these attitudes and beliefs are not expected to change.
WHAT IS MISSING FROM CURRENT RESEARCH?

Future research should focus on detecting health effects that are caused by or associated with e-cigarette use (Born et al., 2015). Although research shows that many carcinogens are inhaled by vapers, existing research does not offer conclusive evidence about specific health effects of these exposures. Researchers, who have studied sentiments, information (such as advertisement, flavors, smoking cessation), and opinions using Twitter data, agree that Twitter data should be collected in order to understand whether Twitter users report any health effects from using e-cigarettes (Cole-Lewis et al., 2015; Godea et al., 2015; Huang et al., 2014; Myslin et al., 2013). Research has shown that health effects caused by smoking are likely to affect the respiratory system (Turner et al., 2015), so this research will assess if Twitter users are reporting respiratory health symptoms including throat irritation, mouth irritation, and coughing.

In addition to examining respiratory health symptoms, the study will examine the public’s sentiments and opinions about e-cigarettes since the e-cigarette prevention campaigns were released in March 2015. The most recent Twitter data study that assessed public opinion about e-cigarettes (Cole-Lewis et al., 2015; Godea et al., 2015; Huang et al., 2014; Myslin et al., 2013) was published before the release of these campaigns. Being able to assess current sentiments and information reported is important to discover trends in knowledge or determine if there has been a shift in public opinion about e-cigarettes (Murphy et al., 2014). Lastly, it is beneficial to determine who is talking about e-cigarettes on Twitter to inform public health professionals that are delivering interventions over social media (Pechmann et al., 2015). In some cases, smoking cessation interventions have been proven to be effective when delivered over social media, so it is valuable to provide source characteristics information to professionals that are creating and delivering interventions to specific populations on social media (Prochaska, Pechmann, Kim, & Leonhardt, 2012). In order to create effective interventions, public health professionals need to know their audience, and Twitter studies can provide information about who is actively discussing e-cigarettes on Twitter by measuring source characteristics. Currently the published studies analyzed in this review only measured a limited number of source characteristics, such as individuals and first-person.
SIGNIFICANCE OF FUTURE RESEARCH

Further research on the respiratory health effects of e-cigarette use should be conducted so that e-cigarette users and potential users are able to make informed health decisions about their choice to use them. Examples of reported respiratory health effects of short-term e-cigarette use are irritated mouth, irritated throat, and dry cough (Farsalinos et al., 2015). Farsalinos et al. (2014) conducted an online survey of more than 19,000 current e-cigarette users in which current smokers, past smokers, and never smokers of combustible cigarette were invited to participate in the survey. Less than 1% of the sample had never used combustible cigarettes, and among this group the most common health effects reported were irritation of the mouth and throat. More research on both short-term and long-term respiratory effects is critical for public health professionals and the general public to know if e-cigarettes are a threat to human health.

Determining current sentiments and beliefs of the public will help public health professionals to design advertisement and interventions that address the current public opinion climate. Assessing current opinion will also provide insight into whether sentiments and opinions have changed since e-cigarette health campaigns have been released.

The primary purpose of the proposed research study is to understand current sentiments around e-cigarette use and information that is being discussed by Twitter users about e-cigarettes. I will code tweets into different categories based on the content posted. The categories will include sentiments (which will be measured by tone), cessation, addiction, stigma, second-hand smoke, health, and versatility. Categories will also include information about smoking e-cigarettes, which will be measured through mention of health effects, policy, advertisements, e-liquid flavors, illicit use, and nicotine. The second objective for my research is to examine what type of Twitter users are discussing e-cigarettes. I will examine this through source characteristics which will give information on who posted the tweet, such as an Individual, Organization, Twitter-Verified Celebrity, or a Potential Robot Account. Examining source characteristics, sentiments, and information will allow me to not only report who is talking about e-cigarettes on Twitter, but also report the specific subjects they are talking about. Measuring sentiments will allow me to examine whether sentiments have changed since the last studies were conducted in 2014. Lastly, I will report the percent of respiratory health effects in the collected Twitter data. Evidence of Twitter users reporting
short-term respiratory health effects associated with e-cigarette use will give researchers insight into whether e-cigarette users are reporting health effects publicly through Twitter.
CHAPTER 3

METHODS

In this study I classified tweets about e-cigarettes into multiple categories based on sentiments, source characteristics, and message characteristics. I coded a random sample of tweets into each category to determine the prevalence of sentiments and health-related information discussed on Twitter about e-cigarettes.

SAMPLE

In this study, I analyzed data from Twitter users (tweets) living in all regions of the United States. In total, there were 193,051 tweets collected between October 2015 and February 2016 that mentioned e-cigarettes. I generated a random sample of 1,000 tweets using a program called Stat Trek, then coded and analyzed the random sample (Stat Trek, 2016). Information about the sources of the tweets was collected from available information on Twitter profiles to determine qualitative characteristics of the sample. Characteristics that were collected include whether the participant is an Individual, Organization, and or a Twitter-Verified Celebrity. Further, if the Twitter user was coded as Individual, information from their profile was also used to determine if they reported being a child, student, parent, spiritual, political, doctor, journalist, health practitioner, or a potential robot account. I coded a profile as Potential Robot Account based on three criteria: (a) if the tweet was identical to another known robot, (b) identical tweets were posted within 30 minutes of each other, and (c) the follow-to-follower ratio (Chu, Giannvecchio, Wang, & Jajodia, 2012). If the Twitter account was coded as Organization, information from their profile was used to determine if they reported being a business entity, health advocacy group, government entity, news/media source, educational institution, health information provider, healthcare delivery organization, or a non-health advocacy organization. The Individual and Organization categories are mutually exclusive, but each of the categories within them are not.
PROCEDURE

This study is a cross-sectional study that was conducted with publicly available tweets collected between October 2015 and February 2016. Because this study uses data from a publicly available social media platform, and does not include or interact with human subjects, Institutional Review Board (IRB) approval was waived for the study.

In order to collect tweets about e-cigarette use, I used software called the Social Media Analytic and Research Testbed (SMART) dashboard (Tsou, Jung, et al., 2015). The SMART dashboard integrates a social media application program interface (API) and geographical information system (GIS) technology to allow its user to visualize geographic trends about specific topics (Tsou, Peddecord, et al., 2015). The specific topics of interest are defined by keywords. The keywords are inputted into the SMART dashboard, and the Twitter data are filtered and collected with a Twitter Search API once a day. The Twitter data are then stored in a database and displayed on a dashboard that creates visuals of the data collected.

I selected strategic keywords to collect tweets about e-cigarettes and e-cigarette use. The keywords I used to collect relevant tweets were modeled off of previous research studies that effectively examined sentiments and information among Twitter users (Cole-Lewis et al., 2015; Huang et al., 2014). The key words used to collect data were: Vaping, Vape, Vaper, Vapers, Vapin, Vaped, Evape, Vaporing, e-cig, ecig, e-pen, epen, e-juice, ejuice, e-liquid, or eliquid. If tweets included any of the above listed keywords, they were determined to have met criteria to be included in the study. The API yielded a total of 193,051 tweets. I conducted content analysis for a subset of 1000 randomly selected tweets. These tweets were used for analysis in the study.

All tweets that met the specified criteria were collected from the universe of tweets between the days of October 28, 2015, and February 6, 2016, a total of 102 days. The only further exclusion criteria were that all tweets in the sample must be written in English, and retweets were not included in the sample. Of the 193,051 tweets that met the study criteria, first I took a random sample of 1,000 tweets and coded them. Later these codes were used for data analysis. I manually coded the randomly selected tweets based on source characteristics and content of the tweet. I conducted coding in accordance with the E-cigarette Tweet Codebook that was developed for this study. After coding, I calculated percentages for the
message and source characteristics of the random sample. In addition to percent, I calculated a chi-square test of independence to examine the relation between Individuals, Organizations, and positive sentiments. More specifically, I examined the relation between Individuals and their likelihood of posting positive sentiments, and compared this to Organizations and their likelihood of posting positive sentiments about e-cigarettes.

**INSTRUMENTATION**

In order to extract the data of interest from tweets, I developed an E-Cigarette Tweet Codebook (see Table 1). The code definitions and categories used in the codebook were adapted from various studies that have measured sentiments and content of tweets, or measured sentiments and knowledge of e-cigarette users (Cole-Lewis et al., 2015; Coleman et al., 2015).

The codebook classified measures of interest by source characteristics and tweet information content. First, the source characteristics specified information on the type of profile that wrote the tweet. The source characteristic codes included Individual, Organization, and Twitter-Verified Celebrities. Twitter-Verified Celebrity is not a mutually exclusive category. For example, a Twitter user can be coded as a Twitter-Verified Celebrity as well as an Organization. The Individual and Organization codes were mutually exclusive, and were broken down further into additional categories. Individuals were classified as any of the following: (a) Parent, (b) Child, (c) Spiritual, (d) Political Figure, (e) Journalist, (f) Doctor, (g) Health Practitioner or (h) Potential Robot Account. Potential Robot Accounts are Twitter accounts that were created by humans, but then use automated features that post tweets using content from Twitter APIs (Chu et al., 2012). I determined if an account should be coded as a Potential Robot Account by using three criteria: (a) the timing of tweets (if identical tweets were posted within 30 minutes of each other), (b) known spam posts, and (c) profile properties such as follower-to-following ratio (Chu et al., 2012). The Potential Robot Accounts code was added to the codebook after I began coding and began to suspect that there was a pattern of robot accounts in the sample. I then re-coded the tweets that had not been examined for robot account behavior. The Political Figure code categorized users as Democrat, Republican, Independent or other. The Doctor code included physicians and dentists. The Health Practitioner code included the following: (a) epidemiologists, (b) health
educators, (c) nurses, (d) psychologists, (e) addiction specialists, and (f) social workers. Organizations were categorized as Business Entity, Health Advocacy Group, Government Entity, News/Media Source, Educational Institution, Health Information Provider, Healthcare Delivery Organization, or Non-Health Advocacy Group. The codes within the Individual and Organization categories were not mutually exclusive. For example, a profile that is coded as an Individual, could be coded as Parent, Spiritual, and Doctor. The Twitter users are coded by the name of their profile, the information provided in their profile, and the information provided in the randomly selected tweet. Detailed information about inclusion and exclusion criteria for source characteristics can be found in Table 1.

The second aspect that this study analyzed was the content of the tweet. Sentiment and information reported in the tweet were the two types of content measured. Both of these served as larger categories from which many smaller sections were created to provide more detailed information about the tweet. The first category in Sentiments was Tone. When analyzing tone, the aspects of interest were whether the tweet conveyed Positive, Negative, Neutral, Ambiguous, or Other sentiments about the use of e-cigarettes for any reason. The Other category was developed to encompass any tweets that were incomprehensible. Other was specifically included because robot accounts often post tweets that are nonsense (Twitter, 2014). Because there was potential for robot accounts to be included in the study sample, an Other category was needed. The next categories of sentiments include specific commonly held ideas by the general public about e-cigarettes. The ideas in this category were derived from various research studies that surveyed e-cigarette users about their thoughts concerning e-cigarette use (Adkison et al., 2013; Choi & Forster, 2013; Coleman et al., 2015; Dawkins et al., 2013; Farsalinos et al., 2014; Goniewicz & Lingas, 2013; Pearson, Richardson, Niaura, Vallone, & Abrams, 2012; Pepper et al., 2014; Richardson, Pearson, Xiao, Stalgaitis, & Vallone, 2014; Zhu et al., 2014). The specific sentiment codes about e-cigarettes were as follows: (a) effectiveness or ineffectiveness in e-cigarette use with smoking cessation, (b) e-cigarettes being more or less addictive than traditional cigarettes, (c) stigma or lack of stigma regarding e-cigarettes, (d) second-hand smoke that is caused or prevented by e-cigarette use, (e) the potential harm or benefits of e-cigarettes on health, (f) the ability or inability for e-cigarette users to vape where they choose, and (g) the dissatisfaction that vapers feel regarding e-cigarettes replacing traditional cigarettes. Lastly,
the objective information in the tweets was analyzed under Information Reported. The subjects coded in Information Reported were: (a) E-Cigarette Policy Information (b) E-Cigarette Advertisements, (c) E-Liquid Flavors, (d) Financial Burden Related to E-Cigarette use, and (e) Illicit Substance Use with an E-Cigarette Device. As an added measure, if tweets mentioned any negative respiratory health effects, such as coughing, sore throat, or lung health, it was coded in a respiratory health category. More detailed information on the coding criteria for these sub-categories can be found in Table 1.

In order to determine which tweets would be used in the random sample of 1,000 tweets that were used for data analysis, I used an online random number generator called Stat Trek (StatTrek, 2016). There were no duplicated numbers in the random sample. I used Microsoft Excel to import the full list of original tweets and assign an identification number to each tweet (identification numbers ranged from 1 to 193,051). Stat Trek generated a list of 1,000 random numbers, to which each of the randomly generated numbers corresponded to an original tweet, which were used for data analysis.

**DATA ANALYSIS**

The primary study objective was to determine the public’s sentiments, knowledge, and information reported regarding e-cigarettes. The measures that were calculated were frequency distributions of the source characteristics and message characteristics of the tweets. In each category of source characteristics, percentage was calculated by dividing the number of times the code occurred by the total number of tweets in the overarching category. For example, the proportion of parents in the sample was derived by dividing the number of times the Parent code occurred by the total number of tweets that were coded Individual.

Frequency proportions gathered from source characteristics will give insight into the percent of different entities that are involved in e-cigarette conversations on Twitter. Second, percent was calculated for each code in the message characteristics section by dividing the number of times the code occurred by the total number of tweets in the random sample.
Table 1. The E-Cigarette Codebook With Definitions of Each Coding Theme in the Sample.

<table>
<thead>
<tr>
<th>Element</th>
<th>Coded Theme</th>
<th>Code</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Characteristic</td>
<td>Individual No=0, Yes=1</td>
<td></td>
<td>Profile represents a single person, not an organization or cause.</td>
</tr>
<tr>
<td>Source Characteristic</td>
<td>Organization No=0, Yes=1</td>
<td></td>
<td>Profile represents an organization, not a single person.</td>
</tr>
<tr>
<td>Source Characteristic</td>
<td>Twitter-Verified Celebrity No=0, Yes=1</td>
<td></td>
<td>This profile is any profile that contains a blue check mark.</td>
</tr>
<tr>
<td>Organization</td>
<td>Business Entity No=0, Yes=1</td>
<td></td>
<td>Profile explicitly mentions a company, franchise, business, store, product, or service. The word “.com” is a qualifier. Excludes Non-profit and Government.</td>
</tr>
<tr>
<td>Organization</td>
<td>Health Advocacy Group No=0, Yes=1</td>
<td></td>
<td>Profile represents an organization that advocates for a specific cause. The word “foundation,” mention of a 501c3, “non-profit,” and “.org” are qualifiers. Excludes Business and Government.</td>
</tr>
<tr>
<td>Organization</td>
<td>Government Entity No=0, Yes=1</td>
<td></td>
<td>Profile represents city, state, or federal government agency. “.gov” is a qualifier. Excludes business and non-profits.</td>
</tr>
<tr>
<td>Organization</td>
<td>News/Media Source No=0, Yes=1</td>
<td></td>
<td>Profile represents a news organization or type of media outlet. Media can be television, print, or digital. “Network” or “providing news” is sufficient. Not mutually exclusive from any other Org category.</td>
</tr>
<tr>
<td>Organization</td>
<td>Educational Institution No=0, Yes=1</td>
<td></td>
<td>Profile cites the organization’s purpose as a form of health education, promotion, advocacy, communication, or new updates. Not mutually exclusive from any other Org category.</td>
</tr>
<tr>
<td>Organization</td>
<td>Health Information Provider No=0, Yes=1</td>
<td></td>
<td>Profile represents a healthcare establishment that is tangible. This can also be coded as Business. Not mutually exclusive from any other Org category.</td>
</tr>
<tr>
<td>Organization</td>
<td>Non-Health Advocacy No=0, Yes=1</td>
<td></td>
<td>Profile represents an advocacy group that is non-health related. Not mutually exclusive from any other Org category.</td>
</tr>
<tr>
<td>Individual</td>
<td>Mother or Father No=0, Yes=1</td>
<td></td>
<td>Any mention on user profile or in tweet of being a mother or a father.</td>
</tr>
<tr>
<td>Individual</td>
<td>Son or Daughter No=0, Yes=1</td>
<td></td>
<td>Any mention on user profile of being a son or daughter, or a child (e.g., someone younger than 18 years).</td>
</tr>
<tr>
<td>Individual</td>
<td>Student No=0, Yes=1</td>
<td></td>
<td>Any mention of the user being a student.</td>
</tr>
<tr>
<td>Individual</td>
<td>First-Person No=0, Yes=1</td>
<td></td>
<td>The tweet is written in first-person. The user mentions a personal experience with e-cigarettes.</td>
</tr>
</tbody>
</table>

(table continues)
Table 1. (continued)

<table>
<thead>
<tr>
<th>Individual</th>
<th>Spiritual No=0, Yes=1</th>
<th>Any mention on user profile or tweet of religion, scripture, prayer, being spiritual, or God.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>Political Persuasion No=0, Yes=1</td>
<td>Any mention on the user profile or tweet of political affiliation or politics.</td>
</tr>
<tr>
<td>Individual</td>
<td>Journalist No=0, Yes=1</td>
<td>Any mention on the user profile or tweet of a news organization or to being a press member/journalist.</td>
</tr>
<tr>
<td>Individual</td>
<td>Doctor No=0, Yes=1</td>
<td>Any mention on the user profile or tweet of being a medical doctor, M.D., D.O., medical school resident, dentist, or dental professional.</td>
</tr>
<tr>
<td>Individual</td>
<td>Health Practitioner No=0, Yes=1</td>
<td>Any mention on the user profile of being an Epidemiologist (not a doctor/dentist) Health Educator, Nurse, Psychologists, Addiction Specialists, and Social Workers.</td>
</tr>
<tr>
<td>Individual</td>
<td>Potential Robot Account No=0, Yes=1</td>
<td>Accounts that appear to be fake/computerized that are primarily promoting e-cigarette products (or other products); most accounts are disguised to appear as Everyday Person.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sentiments</th>
<th>Tone of Tweet</th>
<th>Categorize the tone of the tweet into neutral (e.g., does not approve or disapprove of e-cig use), negative (e.g., disapproves of e-cig use for any reason), positive (e.g., approves e-cig use for any reason), ambiguous (e.g., contains approving and disapproving information), other (e.g., those tweets that did not fit into any other category).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentiments</td>
<td>Effectiveness in Smoking Cessation No=0, Yes=1</td>
<td>The tweet mentions something about e-cigarettes aiding the user in cessation of smoking traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>In-Effectiveness in Smoking Cessation No=0, Yes=1</td>
<td>The tweet mentions something about the ineffectiveness of e-cigarettes in their attempt to quit smoking traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Equal or More Addiction No=0, Yes=1</td>
<td>The tweet mentions e-cigarettes being as addicting or more addicting than smoking traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Less Addiction No=0, Yes=1</td>
<td>The tweet mentions something about the use of e-cigarettes being less addicting than smoking traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Carry Stigma No=0, Yes=1</td>
<td>The tweet mentions something about the use of e-cigarettes carrying a stigma.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Lack Stigma No=0, Yes=1</td>
<td>The tweet mentions something about the use of e-cigarettes having less or no stigma as compared to smoking traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Contribute to Second-Hand Smoke No=0, Yes=1</td>
<td>The tweet mentions something about the use of e-cigarettes contributing to the exposure of second-hand smoke.</td>
</tr>
</tbody>
</table>

(table continues)
Table 1. (continued)

<table>
<thead>
<tr>
<th>Sentiments</th>
<th>Reduce Second-Hand Smoke</th>
<th>The tweet mentions something about the use of e-cigarettes preventing the exposure of second-hand smoke.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentiments</td>
<td>Harmful</td>
<td>The tweet mentions something about e-cigarettes being harmful to the body.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Harmless</td>
<td>The tweet mentions something about e-cigarettes being harmless, or less harmful than traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Versatility</td>
<td>The tweet mentions something about e-cigarettes because being used in most/many/all settings as compared to traditional cigarettes.</td>
</tr>
<tr>
<td>Sentiments</td>
<td>Restricted Use</td>
<td>The tweet mentions something about e-cigarette use being restricted in one or more settings that are also restrictive to traditional cigarettes.</td>
</tr>
<tr>
<td>Information</td>
<td>Policies</td>
<td>The tweet mentions something about a local, state, or federal policy that concerns e-cigarette usage or purchase.</td>
</tr>
<tr>
<td>Information</td>
<td>Advertisement</td>
<td>The tweet is promoting or advertising a specific brand or type of e-cigarette product.</td>
</tr>
<tr>
<td>Information</td>
<td>E-Liquid Flavor</td>
<td>The tweet mentions something about e-cigarette liquid flavors (e.g., vaping juice, e-liquid, e-flavor).</td>
</tr>
<tr>
<td>Information</td>
<td>Illicit Use</td>
<td>The tweet mentions something about e-cigarettes being used to smoke illicit substances (e.g., marijuana).</td>
</tr>
<tr>
<td>Information</td>
<td>Nicotine</td>
<td>The tweet mentions something about the use of nicotine with e-cigarettes.</td>
</tr>
<tr>
<td>Information</td>
<td>Financial</td>
<td>The tweet mentions something about e-cigarettes being a financial burden.</td>
</tr>
<tr>
<td>Information</td>
<td>Respiratory Health Effects</td>
<td>The tweet mentions something about a negative respiratory health effect and the use of e-cigarettes (e.g., throat irritation, mouth irritation, trouble breathing, coughing while using e-cigarettes).</td>
</tr>
</tbody>
</table>

For example, positive sentiments were calculated by dividing all tweets that were coded ‘Positive Sentiment’ by all tweets in the sample. Percent of and specific sentiments regarding e-cigarette use will provide information about the public’s current ideas concerning vaping. The percent calculation for the Information section was derived in two ways. First, a proportion was found for the overall information code by dividing the number of times that subject was mentioned by the total number of tweets in the random sample. For example, the total number of times that cessation was mentioned was divided by the total number of tweets in the sample. Second, within each subject, a percent was derived from the tweets that
confirmed or rejected the subject by dividing the total number of tweets confirming the subject by the total number of tweets that discussed the subject. For example, the total number of tweets that confirmed e-cigarettes as an effective method to smoking cessation was divided by the total number of tweets that mentioned e-cigarettes as a smoking cessation tool. This measure is effective in providing information about whether people are more supportive or more opposed to specific subjects regarding e-cigarettes.

The secondary measure was percent of Twitter users that reported short-term respiratory health effects associated with e-cigarette use. As defined in the codebook, respiratory effects were measured as any mention about a negative respiratory health effect from the use of e-cigarettes (e.g., coughing, throat/lung/mouth irritation, etc.). The reported respiratory health effects percent was derived from the Negative Respiratory Effect code and was calculated by the number of times respiratory health effects were mentioned divided by the total number of tweets in the random sample.

A chi-square test of independence was calculated in order to examine the relationship between organizations, individuals, and positive sentiments. The chi-square statistic tested for the presence of a statistical relationship between type of tweeter (organization versus individual) and the proportion of positive sentiments.

Lastly, because there was potential for robot accounts to be included in the sample, I calculated percentages of source characteristics and tone without potential robot accounts in the sample. This additional analysis allowed me to compare the sample percentages with robot accounts and without robot accounts, to determine if robot accounts are affecting the overall results. Because robot accounts may not be tweeting about real human experiences, these numbers may lend additional insight into the true source characteristics and sentiments of e-cigarettes discussions on Twitter.

To import and code information from tweets, I used Microsoft Excel: mac 2011. In order to analyze the data set, I imported data from Microsoft Excel into IBM SPSS Statistics Grad Pack Base 23. With SPSS, I took percentages for each category and sub-category that were measured through the E-Cigarette Tweet Codebook.
CHAPTER 4

RESULTS

In this section I will summarize the frequencies and percentages that I calculated in each theme category. I will highlight the trends found regarding source characteristics and message characteristics including sentiments, tone, and information.

SOURCE CHARACTERISTICS

The vast majority of tweets in the random sample were identified as individual accounts (89.6%; n=896) and the remainder of tweets were classified as accounts owned by organizations (10.4%; n=104). In the Individual category, a majority of individual senders were potential robot accounts (71.5%; n=641). The next most prevalent sources of tweets were those that identified themselves as Child (9.9%; n=89), those that identified themselves as Student (3.6%; n=33), and those that identified themselves as Journalist (1.9%; n=17). A full description of all proportions are in Table 2. Because each source can be classified into one or more category, percentages in Table 2 do not equal 100%.

Table 2. Descriptive Statistics for Twitter Accounts Coded as Individual. (N=896)

<table>
<thead>
<tr>
<th>The most common Individuals in the sample were:</th>
<th>Observations</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential Robot Account</td>
<td>(n=641)</td>
<td>71.5%</td>
</tr>
<tr>
<td>Child</td>
<td>(n=89)</td>
<td>9.9%</td>
</tr>
<tr>
<td>Student</td>
<td>(n=33)</td>
<td>3.6%</td>
</tr>
<tr>
<td>Journalist</td>
<td>(n=17)</td>
<td>1.9%</td>
</tr>
<tr>
<td>Political Persuasion</td>
<td>(n=14)</td>
<td>1.5%</td>
</tr>
<tr>
<td>Parent</td>
<td>(n=11)</td>
<td>1.2%</td>
</tr>
<tr>
<td>Spiritual</td>
<td>(n=11)</td>
<td>1.2%</td>
</tr>
<tr>
<td>Health Practitioner</td>
<td>(n=4)</td>
<td>0.4%</td>
</tr>
<tr>
<td>Twitter-Verified Celebrity</td>
<td>(n=4)</td>
<td>0.4%</td>
</tr>
<tr>
<td>Doctor</td>
<td>(n=1)</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

Note: The sum of the Percent column adds up to more than 100% because categories were not mutually exclusive.
In the organization category the vast majority of sources in the sample were Business Entities (75.9%; n=79). The second most prevalent organizations were News/Media Sources (21.2%; n=22), and the third most prevalent was Health Advocacy Groups (8.6%; n=9). The two least prevalent sources of tweets were Government Entities (0.9%; n=1) and Educational Institutions (0.9%; n=1). Because each source can be classified into one or more category, percentages in Table 3 do not equal 100%.

Table 3. Descriptive Statistics for Twitter Accounts Coded as Organization. (N=104)

<table>
<thead>
<tr>
<th>The most common Organizations in the sample were:</th>
<th>Observations</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Entity</td>
<td>(n=79)</td>
<td>75.9%</td>
</tr>
<tr>
<td>News/Media Source</td>
<td>(n=22)</td>
<td>21.2%</td>
</tr>
<tr>
<td>Health Advocacy Group</td>
<td>(n=9)</td>
<td>8.6%</td>
</tr>
<tr>
<td>Non-Health Advocacy Group</td>
<td>(n=6)</td>
<td>5.8%</td>
</tr>
<tr>
<td>Health Information Provider</td>
<td>(n=5)</td>
<td>4.8%</td>
</tr>
<tr>
<td>Healthcare Delivery</td>
<td>(n=2)</td>
<td>1.9%</td>
</tr>
<tr>
<td>Government Entity</td>
<td>(n=1)</td>
<td>0.9%</td>
</tr>
<tr>
<td>Educational Institution</td>
<td>(n=1)</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

Note: The sum of Percent adds up to more than 100% because categories were not mutually exclusive.

**MESSAGE CHARACTERISTICS- TONE**

Each tweet in the sample contained Positive, Negative, Neutral, Ambiguous, or Other tone. In the overall sample, the tone was overwhelmingly positive with 68.1% (n=681) of tweets containing supportive (positive) tone about e-cigarette use (see Table 4). The percent of overall tweets that contained an unsupportive (negative) tone about e-cigarette use was 15.2% (n=152), and those that contained a neutral tone were 11.6% (n=116) of the sample. Lastly, 3.6% (n=36) of the tweets contained ambiguous tone regarding use of e-cigarettes. Among sources coded as Individual, 66.5% (n=596) of tweets had a positive tone, versus sources coded as Organization in which 80.7% (n=84) of tweets had a positive tone. Among profiles coded as Child, 55.1% (n=49) of tweets had a positive tone, 27.0% (n=24) of tweets had a neutral tone, and 14.6% (n=13) had a negative tone. Among Organizations, the majority of tweets contained a positive tone, while 12.6% (n=13) contained a negative tone. Business Entities tweeted with mostly a positive tone about e-cigarettes (94.9%; n=75), and tweeted neutral content 3.8% (n=3) of the time.
Alternatively, News/Media Sources tweeted with a negative tone toward e-cigarette use 54.5% (n=12) of the time, and positively 40.9% (n=9) of the time. More descriptive statistics about positive, negative, neutral, and ambiguous tone can be found in Table 4.

Table 4. Percentage of Tweets That Were Coded Positive, Neutral, Negative, Ambiguous, or Other Tone in the Most Prevalent Source Categories.

<table>
<thead>
<tr>
<th>Source</th>
<th>Percent Positive</th>
<th>Percent Neutral</th>
<th>Percent Negative</th>
<th>Percent Ambiguous</th>
<th>Percent Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (N=1,000)</td>
<td>68.1% (n=681)</td>
<td>11.6% (n=116)</td>
<td>15.2% (n=152)</td>
<td>3.6% (n=36)</td>
<td>1.5% (n=15)</td>
</tr>
<tr>
<td>Individuals (N=896)</td>
<td>66.5% (n=596)</td>
<td>12.5% (n=112)</td>
<td>15.5% (n=139)</td>
<td>4.0% (n=36)</td>
<td>1.5% (n=13)</td>
</tr>
<tr>
<td>Potential Robots (N=642)</td>
<td>69.2% (n=444)</td>
<td>11.2% (n=72)</td>
<td>14.3% (n=92)</td>
<td>4.4% (n=28)</td>
<td>0.9% (n=6)</td>
</tr>
<tr>
<td>First-Person (N=528)</td>
<td>91.7% (n=484)</td>
<td>2.3% (n=12)</td>
<td>1.3% (n=7)</td>
<td>4.5% (n=24)</td>
<td>0.2% (n=1)</td>
</tr>
<tr>
<td>Child (N=89)</td>
<td>55.1% (n=49)</td>
<td>27.0% (n=24)</td>
<td>14.6% (n=13)</td>
<td>3.4% (n=3)</td>
<td>0.0% (n=0)</td>
</tr>
<tr>
<td>Student (N=33)</td>
<td>60.6% (n=20)</td>
<td>18.2% (n=6)</td>
<td>21.2% (n=7)</td>
<td>0.0% (n=0)</td>
<td>0.0% (n=0)</td>
</tr>
<tr>
<td>Parent (N=11)</td>
<td>63.6% (n=7)</td>
<td>0.0% (n=0)</td>
<td>27.3% (n=3)</td>
<td>0.0% (n=0)</td>
<td>9.1% (n=1)</td>
</tr>
<tr>
<td>Organizations (N=104)</td>
<td>80.7% (n=84)</td>
<td>4.9% (n=5)</td>
<td>12.6% (n=13)</td>
<td>1.0% (n=1)</td>
<td>1.0% (n=1)</td>
</tr>
<tr>
<td>Business Entity (N=79)</td>
<td>94.9% (n=75)</td>
<td>3.8% (n=3)</td>
<td>1.3% (n=1)</td>
<td>0.0% (n=0)</td>
<td>0.0% (n=0)</td>
</tr>
<tr>
<td>News/Media (N=22)</td>
<td>40.9% (n=9)</td>
<td>4.5% (n=1)</td>
<td>54.5% (n=12)</td>
<td>0.0% (n=0)</td>
<td>0.0% (n=0)</td>
</tr>
</tbody>
</table>

Note: Each row’s sum is 100%; categories within ‘Individuals’ and ‘Organizations’ are not mutually exclusive.

Lastly, the relation between Organizations, Individuals and Positive Sentiments was examined by conducting a chi-square test of independence. The chi-square analysis indicated that the relation between these variables was significant $X^2 (1, N=680) = 639.09$, p<.01. The significance of the results means that Organizations in the sample are statistically significantly more likely than Individuals to tweet a positive tone about e-cigarettes.

More than half of individuals tweeted about first-person e-cigarette use (58.9%; n=528). These individuals specifically had experience, and tweeted about using e-cigarettes themselves.
MESSAGE CHARACTERISTICS- SENTIMENTS

There were a variety of topics discussed in the sample that addressed commonly held ideas about e-cigarettes. Tweets that reflected commonly held ideas were categorized as confirmed or rejected. For example, Confirmed Cessation means that the sender reported that e-cigarettes are effective in helping cigarette smokers quit smoking, and Rejected Cessation means that the sender reported that e-cigarettes are not effective in helping cigarette smokers quit smoking. The most prevalent subject that Twitter users discussed overall was the stigma associated with e-cigarettes. Overall, 11.7% (n=117) of tweets in the sample discussed the subject of stigma regarding e-cigarette use. Of the 117 tweets that discussed stigma, 70.9% (n=83) of users confirmed that a stigma does exist for people that use e-cigarettes (e.g., “My friend just asked me if I wanted to vape. We had a good run man ill see ya around”) and 29.1% (n=34) of users denied that there was a stigma associated with e-cigarette use (e.g., “life fact: putting ‘I vape’ on ur tinder bio actually gives you hella matches”). The second most prevalent subject discussed was harm that may or may not be caused by the use of e-cigarettes. Overall, 6.7% (n=67) of all tweets discussed harm. Of the 67 tweets that discussed harm, 46.2% (n=31) of them cited that e-cigarettes are harmful to your health (e.g., “my sisters friends vape exploded on her”), and 57.3% (n=36) cited that e-cigarettes are not harmful or significantly less harmful than traditional cigarettes (e.g., “sighs smokes real hard on my lungs its kind of embarrassing i can literally only vape bc its so much more gentle”). Additional instances in which the harm code was used was when tweets reported e-cigarette users getting shocked by their vape, claimed that e-cigarette use would cause lung cancer, and when hot vape liquid would burn the e-cigarette user. In one instance, a student’s backpack caught on fire because of the e-cigarette battery and the school building had to be evacuated. The third most prevalent subject discussed was second-hand smoke caused by e-cigarettes. Overall, 5.3% (n=53) of Twitter users mentioned the idea of e-cigarettes and second-hand smoke. Among these tweets 94.3% (n=50) confirmed the existence of second hand smoke caused by e-cigarettes (e.g., “I dont vape but I do like secondhand vaping”), and 5.7% (n=3) rejected the idea that e-cigarettes cause second hand smoke (e.g., “Why do vapers have to isolate themselves in smoking areas? Like, I'm sorry that my water vapor which smells like hi-chews offends you ???”). Further descriptive statistics about other
message characteristics, including cessation, versatility of e-cigarettes, and e-cigarette addiction, can be found in Table 5.

Table 5. Percentage of Tweets That Discussed Commonly Held Sentiments About E-Cigarettes. (N=1,000)

<table>
<thead>
<tr>
<th>Sentiments</th>
<th>Overall</th>
<th>Confirmed Percentage</th>
<th>Rejected Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stigma (N=117)</td>
<td>11.7%</td>
<td>(n=83) 70.9%</td>
<td>(n=34) 29.1%</td>
</tr>
<tr>
<td>Harmfulness (N=67)</td>
<td>6.7%</td>
<td>(n=31) 46.2%</td>
<td>(n=36) 57.3%</td>
</tr>
<tr>
<td>Second-Hand Smoke (N=53)</td>
<td>5.3%</td>
<td>(n=50) 94.3%</td>
<td>(n=36) 5.7%</td>
</tr>
<tr>
<td>Cessation (N=36)</td>
<td>3.6%</td>
<td>(n=28) 77.8%</td>
<td>(n=8) 22.2%</td>
</tr>
<tr>
<td>Versatility (N=30)</td>
<td>3.0%</td>
<td>(n=22) 73.3%</td>
<td>(n=8) 26.7%</td>
</tr>
<tr>
<td>Addiction (N=4)</td>
<td>0.4%</td>
<td>(n=2) 50%</td>
<td>(n=2) 50%</td>
</tr>
</tbody>
</table>

Note: Confirmed Percentage is the percentage of Twitter users that mentioned a situation in which the commonly held belief exists. For example, Confirmed Harmfulness is a tweet that mentions an instance where e-cigarettes were harmful to the user. Rejected Percentage is the percentage of Twitter users that mentioned a situation in which the commonly held belief did not exist. For example, Rejected Harmfulness is a tweet that mentions an instance where e-cigarettes were not harmful, or less harmful than traditional cigarettes.

**MESSAGE CHARACTERISTICS- INFORMATION**

There were a variety of topics regarding e-cigarettes mentioned in tweets. The most prevalent topics were advertisements (7.2%; n=72) and mentions of specific flavors of e-liquids (7.0%; n=70). The third most prevalent subject was illicit use of e-cigarettes (4.8%; n=48). All illicit mentions of e-cigarettes discussed marijuana use. Further statistics regarding financial burden of e-cigarettes for e-cigarette users, e-cigarette policies mentioned, and tweets discussing nicotine in e-cigarettes can be found in Table 6.
Table 6. Percentage of Tweets Discussing Specific Subjects About E-Cigarettes.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Observations (n)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertisement</td>
<td>72</td>
<td>7.2%</td>
</tr>
<tr>
<td>E-Liquid Flavors</td>
<td>70</td>
<td>7.0%</td>
</tr>
<tr>
<td>Illicit Use</td>
<td>48</td>
<td>4.8%</td>
</tr>
<tr>
<td>Financial</td>
<td>35</td>
<td>3.5%</td>
</tr>
<tr>
<td>Policy</td>
<td>21</td>
<td>2.1%</td>
</tr>
<tr>
<td>Nicotine</td>
<td>16</td>
<td>1.6%</td>
</tr>
</tbody>
</table>

Note: The sum of Percent adds up to less than 100% because every tweet in the sample was not required to fit into one of the categories. Categories were not mutually exclusive.

I also sought to gather information regarding Twitter users that reported negative respiratory effects as a result of e-cigarette use. The results showed that none of the 1,000 tweets discussed negative respiratory health effects.

**Sentiments and Sources Without Potential Robot Accounts**

Because of the high percentage of Potential Robot Accounts, and the possibility that they may not be humans tweeting, I conducted additional analyses that excluded all potential robot accounts to determine how the percentages differed. Without potential robot accounts the overall sample included 359 cases. Individual accounts made up 71.0% (n=255) of the sample, and organizations made up 28.9% (n=104) of the sample. Table 7 shows that the most prevalent Individual accounts were instances coded as First-Person experiences with e-cigarettes (43.9%, n=112) and Students (9.4%, n=24).

Table 7. Descriptive Statistics for Twitter Accounts Coded as Individual, Among Cases That are Not Coded as Potential Robot Accounts. (N=255)

<table>
<thead>
<tr>
<th>The most common Individuals in the sample were:</th>
<th>Observations (n)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals</td>
<td>255</td>
<td>71.0%</td>
</tr>
<tr>
<td>First-Person</td>
<td>112</td>
<td>43.9%</td>
</tr>
<tr>
<td>Student</td>
<td>24</td>
<td>9.4%</td>
</tr>
<tr>
<td>Child</td>
<td>22</td>
<td>8.6%</td>
</tr>
<tr>
<td>Parent</td>
<td>10</td>
<td>3.9%</td>
</tr>
</tbody>
</table>

Note: The sum of the Percent column adds up to more than 100% because categories were not mutually exclusive.

I also calculated percentages of tone among cases that were not coded Potential Robot Accounts. Overall, more than half of sentiments were positive (66.3%, n=238) and 16.7%
(n=60) were negative. Among Individuals, more than half of users also tweeted positive sentiments (60.4%, n=155), and negative sentiments were slightly higher at 18.5% (n=47). Those Individuals that tweeted about first-hand experience with e-cigarettes indicated overwhelmingly positive sentiments (92.9%, n=105) and only 1.8% negative sentiments. Further information about neutral, ambiguous, and other sentiments can be found in Table 8.

**Table 8. Percentage of Tweets That Were Coded Positive, Neutral, Negative, Ambiguous, or Other Tone Among Cases That Were Not Coded Potential Robot Account.**

<table>
<thead>
<tr>
<th>Source</th>
<th>Percent Positive</th>
<th>Percent Neutral</th>
<th>Percent Negative</th>
<th>Percent Ambiguous</th>
<th>Percent Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (N=359)</td>
<td>66.3% (n=238)</td>
<td>12.3% (n=44)</td>
<td>16.7% (n=60)</td>
<td>2.2% (n=8)</td>
<td>2.5% (n=9)</td>
</tr>
<tr>
<td>Individuals (N=255)</td>
<td>60.4% (n=154)</td>
<td>15.7% (n=40)</td>
<td>18.5% (n=47)</td>
<td>2.7% (n=7)</td>
<td>2.7% (n=7)</td>
</tr>
<tr>
<td>First-Person (N=113)</td>
<td>92.9% (n=105)</td>
<td>0.9% (n=1)</td>
<td>1.8% (n=2)</td>
<td>4.4% (n=5)</td>
<td>0.0% (n=0)</td>
</tr>
</tbody>
</table>

Note: Each row’s sum is 100%. The categories within ‘Individuals’ and ‘Organizations’ are not mutually exclusive.

Lastly, because there was a change in Individual positive sentiment when Potential Robot Accounts were removed, I also calculated the relation between Organizations, Individuals, and positive sentiments by conducting a chi-square test of independence with all Potential Robot Accounts removed. It is important to test for a relationship among these variables without Potential Robot Accounts (true individuals) in the sample because public health professionals are interested in understanding human behaviors and sentiments. Potential Robot Accounts may not be tweeting human sentiments, so removing them from the sample might represent a more accurate depiction of human sentiments about e-cigarettes. The chi-square analysis indicated that the relation between organizations, true individuals, and positive sentiments was significant \( \chi^2 (1, N=358) = 341.53, p<.01 \). The significance of the results means that organizations in the sample are more likely than true individuals to tweet a positive tone about e-cigarettes.
CHAPTER 5
DISCUSSION

This study examined the source characteristics and message characteristics of 1,000 randomly selected tweets from Twitter users discussing e-cigarettes. I measured the type of Twitter accounts that were tweeting about e-cigarettes (e.g., Individual accounts versus Organizations), and also what subjects the Twitter users were discussing (e.g., Smoking Cessation, Harm of E-Cigarettes, Second-Hand Smoke). The results indicated that there were more than double the amount of Individuals tweeting about e-cigarettes than Organizations, and that more than half of the Individual accounts were Potential Robot Accounts. Robot accounts are created and managed by a human but rely on a robot to automatically post content intermittently (Chu et al., 2012). Results also indicated that Organizations were statistically more likely than Individuals to post tweets with positive tone regarding e-cigarettes. When Potential Robot Accounts were removed from the study, percentages of positive and negative sentiments were very similar to percentages with the Potential Robot Accounts. I also conducted a chi-square test of independence without Potential Robot Accounts, and found that Organizations remained more likely than Individuals to tweet positive sentiments. Therefore, Potential Robot Accounts did not have a significant influence on the percentage of positive and negative sentiments. When Potential Robot Accounts were removed from the sample for source characteristics, it decreased the percentage of Individuals in the sample (from 89.6% to 71.0%). The most prevalent subjects that Twitter users were discussing about e-cigarettes were regarding Advertisements and specific E-liquid Flavors. Tweets coded as Advertisements included links to websites where e-cigarettes could be purchased, or included promotional marketing for e-cigarettes or e-liquids. Regarding commonly held beliefs, the most prevalent subjects were mentions of the Stigma of e-cigarettes and Harm of e-cigarettes. Although some users mentioned Harm caused by e-
cigarettes, none of the tweets in the sample specifically referenced negative respiratory effects in conjunction with e-cigarette use.

This study collected over 190,000 e-cigarette-related tweets across a 102-day span, which is comparable to a similar e-cigarette study conducted in 2014 that used similar technology to collect relevant tweets (Godea et al., 2015). The study that Godea et al. (2015) conducted is similar in that it also collected over 150,000 tweets about e-cigarette-related tweets, and found that more than half of sentiments about e-cigarettes were positive. The abundance of tweets collected that mention e-cigarettes, provide limited evidence that awareness and discussion of e-cigarettes online remains high.

The current study is also consistent with Cole-Lewis et al. (2015), who found that a majority of accounts in the sample were identified as individual users. The current study found that almost 90% of tweets were sent from individual accounts (versus organizations), and Cole-Lewis et al. (2015) found that more than 85% of e-cigarette-related tweets came from individual accounts. Similarly, Kavuluru and Sabbir (2016) found that regular tweeters (e.g., individuals) made up 90% of their dataset. Cole-Lewis et al. (2015) and Kavuluru and Sabbir (2016) were the only other Twitter studies, to my knowledge, that examined source characteristics of e-cigarette users.

Another prominent feature of this study is that it measured the proportion of Potential Robot Accounts. Cole-Lewis et al. (2015) was also the only study that measured the percent of tweets about e-cigarettes from robots. Overall, Cole-Lewis et al. (2015) found under 10% of tweets were sent by a robot account. In contrast, the current study found that over 60% of all tweets in the sample were sent by Potential Robot Accounts. Upon further inspection of this discrepancy, I found that in one study Cook, Waugh, Abdipanah, Hashemi, and Rahman (2014) declared that there is uncertainty on the prevalence of robot accounts on social media, but because of their ability to automate tweets quickly, they have the ability to influence online conversations. He found that in conversations of politics, robot accounts were prevalent (Cook, et al, 2014). In the same year, Twitter (2014) reported that up to 23 million of its users might be robots (Twitter, 2014). Further, it in one study Edwards, Edwards, Spence, and Shelton (2014) found that humans perceive robot accounts as credible sources of reliable information, so it is important for researchers to consider the influence that potential robot accounts have on the online e-cigarette environment.
Twitter robots are often used for two reasons: (a) for advertisement reasons by businesses and (b) for creative reasons by companies or individuals. When robots are used for advertising, the robot will automatically insert advertisement URLs into tweets and post them. The increased creation of robot accounts to disseminate positive Twitter sentiments about e-cigarettes may be possible because e-cigarettes are a growing economic market. It would be in the best interest of businesses in the e-cigarette market to create robot profiles that disseminate positive information about e-cigarettes. Robots can also be used by companies or individuals for creative reasons. In these instances the robot automatically generates tweets that are similar to other accounts that they are following. Creative robot accounts are used for entertainment and can increase awareness of a specific topic. For example, a popular Twitter robot called Real Human Praise posts tweets that are typical of conservative news organizations. Examining the influence of the robot’s sentiments on the overall sample could assist in determining the reason for a high percentage of Potential Robot Accounts in the sample.

In the current study I examined this further and looked into how Potential robot accounts influenced the percentage of source characteristics and tone in the sample. Upon removing the potential robot accounts from the sample, the percentage of individual accounts declined more than 15% (from 89.6% to 71.0%), but the majority of Twitter users in the sample remained as individual users. Among the overall sample (excluding potential robot accounts), overall tone did not change considerably. Over the five Tone categories, differences before and after robot accounts were removed ranged from 0.7% to 1.8%. Among individuals, there were larger changes in percentages. When potential robot accounts were removed positive sentiments declined by 6.1%, but the majority of individual sentiments were still positive. Even smaller differences were found among first-person accounts when potential robot accounts were removed. These differences in tone changed from a range of 0.1% to 1.2%. Although removing potential robot accounts altered some percentages, it did not significantly alter the distribution of sentiments in the sample.

Based on these results, I compared the feasibility of the two common types of robot accounts, advertisement and creative. Robots using Twitter for e-cigarette advertisement is not consistent with the data that was collected in this study, because none of the potential robot accounts contained advertisements for e-cigarettes. The robot accounts used for
creative reasons may be a more feasible explanation for the high percentage of potential robot accounts in my sample, but because I did not code for this, no conclusions can be made. Further research will be needed to examine this change in percent of potential robot tweets and their influence on the e-cigarette social media environment.

Because tweets from potential robot accounts are expansive, further research should be done on the Twitter robot’s influence on social media. For example, researchers can measure Klout scores, a measure of social media influence that is based on a user’s followers and interactions with other social media users (Klout, 2016). This means that robot tweets may be prevalent, but if their Klout scores are low, they do not reach a large amount of Twitter users so they are not likely to be influential.

The current study found that advertisements and specific e-liquid flavors were the most prevalent subjects discussed among tweets in the sample. These findings are consistent with a study conducted by Rath et al. (2012), which found that increasing prevalence of e-cigarette advertisements and marketing of appealing flavored e-juice is a major contributor to the heightened awareness and use of e-cigarettes among youth and adult populations. The current study found that only 7.2% of tweets were advertising e-cigarettes which is consistent with Myslin et al. (2013) who found that 9% of tweets were advertisements. Three additional studies found different results, however, reporting that e-cigarette marketing was much more prevalent on Twitter. Godea et al. (2015) found that more than 20% of tweets were advertising e-cigarettes, Clark et al. (2016) found that 80% of tweets were advertisements, and Huang et al. (2014) found that almost 90% of tweets were advertising e-cigarettes. The discrepancy may be dependent on the amount of Potential Robot Accounts that were found in the current study’s sample because, among organizations that were tweeting about e-cigarettes, almost half of tweets were e-cigarette Advertisements.

A difference between sentiments tweeted among Organizations and sentiments tweeted among Individual accounts was found. Unlike past studies, this study measured the likelihood of Organizations and Individuals tweeting positive content, and found that Organizations are more likely to tweet positive content than Individuals. When considering the influence of organizations on social media, it is important for researchers and public health professionals to note that organizations are disseminating mostly positive information about e-cigarettes. This is likely because more than three-quarters of organizations in the
sample were Business Entities, many of which were selling e-cigarettes. Other organizations, such as Health Advocacy Groups and Government Entities, contained tweets with negative sentiments about the potential dangers of e-cigarettes, but these were not the majority.

This study reports that most Twitter users in the sample discussed e-cigarettes with a positive tone. These findings are consistent with other Twitter studies that measured positive and negative sentiments regarding e-cigarettes (Godea et al., 2015; Myslin et al., 2013). The current study also found that the percent of neutral tone was comparable to the percent of negative tone about e-cigarettes. Past studies have not measured the percent of neutral tone. The percent of neutral sentiments among Twitter users is an important measure for researchers and health educators to consider because it shows that many people may not be knowledgeable about the health considerations of e-cigarettes, which presents an important opportunity for continued education.

It is important to discuss that although overall sentiments were mostly positive, almost 12% of tweets acknowledged a stigma associated with e-cigarette use. The stigma of e-cigarettes did not concern negative health effects, but rather referenced that the use of e-cigarettes are uncool. Users that referenced stigma discussed instances in which they or their friends were un-approving of e-cigarette use, but it is important to note that the tweets did not discuss the health implications of e-cigarette use.

This study also found that almost three quarters of users, who talked about versatility of e-cigarettes, mentioned instances in which e-cigarette use was permitted in locations where traditional cigarette smoking would not be permitted. The high frequency of tweets about the versatility of e-cigarette smoking is preliminary evidence that suggests that people perceived greater versatility with e-cigarettes than with traditional cigarettes. If public health professionals aim to reduce use of e-cigarettes, then restricting their use in public and private spaces through policy, may be an effective way to inhibit use.

Public health professionals have identified the use of illicit substances with vapes, specifically marijuana, among youth as a health concern (Budney, Sargent, & Lee, 2015). Vapes present a unique concern with illicit substances because marijuana cannot be detected when smoked through a vape, so youth may be able to use the substance discreetly. The current study found that although there is not a large proportion of overall tweets mentioning the use of marijuana in vapes, it was the third most mentioned subject among Twitter users.
Cole-Lewis et al. (2015) also measured illicit use of vapes and found that less than 2% of users were tweeting about marijuana use in 2014, and in the current study, that number more than triples. The increased percent of marijuana-related e-cigarette tweets may be evidence that illicit substance use is becoming more prevalent among Twitter users that tweet about e-cigarettes. Another explanation is that the HDMA dashboard drew an unrepresentative sample of marijuana-related tweets. Because the dashboard only collects a small percent of total public tweets, it is possible that the sample was not representative of the true population of Twitter users in the US.

Twitter users who mentioned Harm due to the use of e-cigarettes were the second most prevalent tweets in the sample. More than half of tweets emphasized that e-cigarettes were not harmful to health or less harmful than traditional cigarettes. More than 45% of these tweets discussed instances in which e-cigarette use is harmful. The tweets that highlighted harm caused by e-cigarettes did not have to do with respiratory effects, as the e-cigarette awareness campaigns warn the public about (CDPH, 2015; Healthy Chicago, 2015; HHS, 2015). Rather, harm codes were discussing instances in which e-cigarette batteries blew up in their users’ pockets, in their face, or in their hands. There were also scenarios in which the e-cigarette device shocked its user. Malfunctions with e-cigarette technology, such as these examples, were not found in the literature as a prominent cause for concern. This research brings these dangers to light, and advises that the safety of e-cigarette technology should be considered in their health assessment.

Future research should collect geographic information from tweets and compare opinions and sentiments based on geographic location in the US. Examining information from different geographic locations in the United States may uncover different sentiments and opinions in different regions of the US. Regional information will help public health professionals create targeted interventions that are specific to the opinions and sentiments of populations of a specific region. The current study had a very small percentage of tweets that provided the geographic location of the sender (0.01%; n=6) (e.g., they were geo-tagged tweets), therefore geographic analysis could not be run with the data set. One obstacle to analyzing geographic information is that Twitter users often do not provide accurate information regarding their location. In order for future studies to gain higher percentages of
validated geographic information, Twitter would need to offer this information to researchers from their database.

This research study is not without limitations. One limitation is that the information that I used to code the source characteristics was limited to the information that the Twitter user provided. For example, if a Twitter user was a high school student under the age of 18 years, but did not indicate those characteristics anywhere on their Twitter profile or in the randomly selected tweet, that source was not coded as Student or as Child. Another limitation is that data collected from Twitter is not representative of populations that do not use Twitter. This means that these study results cannot be generalized to the rest of the total US population. This sample may not be representative of the whole Twitter population either. The SMART dashboard used to collect tweets, can only collect a certain percent of available tweets. Lastly, this research describes the content of people’s sentiments about e-cigarettes, but the research does not give any information about the health effects of e-cigarettes.

There are many strengths to this research study. The first is that a large amount of data were collected across more than 100 days. Although this study is a cross-sectional study, content was collected over a long period of time in order to gain differing views about users and e-cigarettes. Another strength of this study is that the geographic areas in which data were able to be collected was expansive: Data were collected from every region in the US. In comparison to other studies that used Twitter data to assess sentiment or content of tweets regarding e-cigarettes, this study offered a comprehensive view in that it assessed source characteristics, sentiment and content. The current study also included greater description of the subjects. For example, while Cole-Lewis et al. (2015) assessed if smoking cessation was mentioned in tweets, the current study assessed whether user declared that e-cigarettes were effective in assisting with smoking cessation or if users asserted that e-cigarettes were ineffective in assisting with traditional smoking cessation. Lastly, this study adds to the current literature by giving an update on the current sentiment and beliefs that Twitter users report about e-cigarettes.

The current study implies that public health professionals have an opportunity to educate and increase awareness around the potential dangers of e-cigarettes. Although much of the sentiment around e-cigarettes is positive, there is also a sizeable amount of sentiment from this sample that is neutral, which is why public health professionals should focus more
on educating populations that have not formed strong opinions about e-cigarettes. Teaching about the influence of Big Tobacco on e-cigarette promotion and youth-targeted advertising may also be an effective way to draw a connection between e-cigarettes and traditional cigarette smoking. Because this study shows that businesses are very prevalent among e-cigarette advertising and promotion, it could be used as an example of how prevalent advertising is being used to promote the use of e-cigarettes among all populations. Lantz et al. (2000) found that increasing awareness about the involvement of Big Tobacco and their targeted advertisement strategies is an effective way to prevent youth from initiation of e-cigarette use. This tactic is called the social resistance model, and it teaches youth skills that give them the ability to recognize and resist social influences related to initiation of smoking, including advertising tactics (Lantz et al., 2000).

It is highlighted in Healthy People 2020 objectives that prevention of youth and adult initiation of tobacco use is among the top health priorities (Healthy People 2020, 2013). This study is important to this objective because e-cigarette use among youth and adults is growing at a rapid rate, and there is evidence that shows that e-cigarettes may serve as a gateway to traditional cigarette smoking (Morris et al., 2016). Because we still have limited information on the effects of e-cigarettes, it is important for researchers to collect and analyze additional data on the public’s view of e-cigarettes. Further research on e-cigarettes is also critical because currently public health professionals’ voices are divided between those that believe in preventing e-cigarette use and those that believe in promoting e-cigarette use for tobacco harm-reduction. Additional research will allow public health professionals to understand the potential positive and negative effects of e-cigarettes, and work to create a united voice regarding the advantages and dangers when using e-cigarettes. When public health professionals are able to assert a united voice about e-cigarette use, it will allow the public to be more accurately informed about this new technology.

Through this study, I found that businesses are likely using Twitter to market e-cigarettes and promote the positive aspects of e-cigarette use. Also, robot accounts are expansive in e-cigarette conversations on Twitter and the origins, intentions, and influence of these users need to be examined further. It is important for public health professionals to know that the general public currently expresses mostly positive tone toward e-cigarettes, that many people see that they are helpful for smoking cessation, but that many non-users of
e-cigarettes associate a stigma with e-cigarette use. Moving forward, public health professionals will continue to have a critical role in the health education of the public surrounding e-cigarettes, and e-cigarettes’ effect on population health.
REFERENCES


