



Published in final edited form as:

Am J Prev Med. 2019 February ; 56(2 Suppl 1): S65–S75. doi:10.1016/j.amepre.2018.07.039.

Identifying Potential Campaign Themes to Prevent Youth Initiation of E-Cigarettes

Angeline Sangalang, PhD¹, Allyson C. Volinsky, MA², Jiaying Liu, PhD³, Qinghua Yang, PhD⁴, Stella J. Lee, PhD^{5,6}, Laura A. Gibson, PhD², Robert C. Hornik, PhD²

¹Department of Communication, University of Dayton

²Annenberg School for Communication, University of Pennsylvania

³Department of Communication Studies, University of Georgia

⁴Department of Communication Studies, Texas Christian University

⁵Dana-Farber Cancer Institute

⁶T.H. Chan School of Public Health, Harvard University

Abstract

Introduction: Once a target audience and a health behavior of interest are selected for a potential media campaign, the next task is selecting beliefs about the health behavior to serve as the basis for campaign message content. For novel health behaviors, such as the use of emerging tobacco products, limited empirical research on beliefs about these behaviors exists. We applied a multi-method approach to generate potential campaign beliefs for emerging behaviors.

Methods: We conducted three methods in this investigation in order to generate a list of potential testable campaign beliefs, using youth electronic cigarette use as a case study: (1) a search of published and unpublished literature including gathering measures from several national surveys, (2) an online elicitation survey, and (3) unsupervised topic modeling of media texts.

Results: We detail how we employed each method to both generate and prioritize beliefs related to youth e-cigarette use into a final set of 115 beliefs across 23 belief themes.

Discussion: We argue this multi-method approach can provide four utilities when thinking through a health campaign for novel health behaviors: 1) developing an exhaustive and complementary list of beliefs; 2) generating overarching themes and distilling larger themes into more nuanced beliefs; 3) identifying language most relevant to the target population; and 4) prioritizing beliefs for message pilot testing with members of the target audience.

Once a health campaign selects a target behavior, an immediate next step is identifying message content to serve as the basis of the campaign initiative. Theories of behavioral prediction such as the Reasoned Action Approach¹ and the Health Belief Model² underscore

This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Corresponding Author Information: Correspondence regarding this article should be directed to Angeline Sangalang, Ph.D., Assistant Professor, Department of Communication, University of Dayton, 300 College Park Avenue, Dayton, OH 45469 (937-229-2819; asangalang1@udayton.edu).

the importance of beliefs in eventual adoption of a health behavior. Beliefs about similar content can also be grouped into large themes to serve as broader guides for campaign development. Our work provides an approach to identify both specific beliefs and belief themes for pilot testing with a target audience and subsequently future campaign message development. *The Real Cost* campaign employed models of behavioral prediction in formative research to identify several campaign belief themes.³ Beliefs about a focus behavior can be directly targeted by campaign messages, and it is known from past work that identifying key target beliefs can lead to campaign success.⁴ Furthermore, recent evaluative evidence suggests that formative work on *The Real Cost*'s message strategies was instrumental in the campaign's success at shifting youth's beliefs and attitudes related to tobacco use.⁵⁻⁸

However, as future campaign efforts attempt to build on the success of *The Real Cost*, particularly for emerging tobacco products such as electronic cigarettes [e-cigarettes], hookah, and cigarillos, several challenges to campaign topic selection are presented: these products lack a magnitude of empirical research evidence (that exists for traditional tobacco products) and also have evolved much more quickly. We applied a multi-method approach to identify potential target campaign beliefs, focusing on a youth (13–17 year olds) e-cigarette prevention campaign as a case study. Due to the novelty of e-cigarettes at the time, we could not simply identify beliefs from the literature as we did in the original *The Real Cost* campaign targeted at preventing traditional cigarette use. We began an iterative, multi-method approach that helped generate beliefs for message testing. We detail the three methods employed (i.e., review of existing literature and scholarly resources, elicitation surveys, and unsupervised topic modeling [TM] of media databases) to generate a list of potential e-cigarette beliefs for a youth-focused prevention campaign. In the following sections, we describe the procedures, benefits, and challenges of each method.

Three Methods for Generating Potential Campaign Themes

Campaign strategies and message development depend on the identification of specific beliefs and belief themes. Holding specific beliefs about a health behavior can predict readiness to perform that behavior.¹⁻² Therefore, a natural first step in campaign development is identifying specific beliefs that can later be utilized in message creation. Several specific beliefs may be related to a belief theme. For example, the notions that cigarette use discolors one's teeth and that cigarette use leads to the formation of wrinkled skin are two distinct beliefs about cigarette use, though both fall under the broader theme of the impact of cigarette use on appearance. Aggregating specific beliefs that predict behavior into themes can be used in messages that desire to incorporate a series of different but related consequences. Conversely, campaign developers may have a target theme in mind but may need guidance on the specific messages (beliefs) that underpin that theme. Therefore, identifying themes and beliefs is somewhat reciprocal.

In the present investigation, we initially wanted to replicate the process of generating specific beliefs and belief themes that was used for *The Real Cost* campaign, by doing an exhaustive literature review. It quickly became apparent that such a literature review would not generate enough belief themes and specific beliefs because so few empirical studies had

explored perceptions of e-cigarettes, particularly among youth. To reach the ultimate goal of generating a thorough list of beliefs, we began to use supplementary methods simultaneously (an elicitation survey and topic modeling) to iteratively generate beliefs.

We documented the beliefs we found in a master database. Any unique belief was added, even if it only had slight differences in wording (e.g., “E-cigarettes are addicting” and “could get addicted” are arguably the same belief with the same specificity but were initially entered as separate beliefs). In total, 564 specific beliefs were collected through this process. In order to distill this large list into a smaller final list for message testing, beliefs that were similar in content (such as “addicted” vs. “addicting”) were collapsed into single beliefs. Then this slightly smaller set of beliefs were grouped by topic similarity together into belief themes (e.g., “yellow teeth” and “wrinkles” into cosmetic effects/appearance theme). Then, beliefs and themes were prioritized as important based on how often they appeared across multiple methods. This process became somewhat iterative as draft themes were created, collapsed, expanded, or revised as new beliefs were identified or specified. Though imperfect, we found this process to be helpful in generating a testable list of beliefs. It is our hope that outlining this procedure will help other researchers exploring emergent health behaviors. We begin by describing the procedures for each of the three methods, as well as their benefits and challenges.

Method 1: Review of Existing Scholarship

A natural first step in campaign development is examining extant work on the focus behavior both in peer-reviewed journals and recent academic conferences. In formative research, past work can help indicate the relationship between specific beliefs and behaviors, thus providing guidance for identifying promising beliefs or themes that might lead to future behavior change.

In the present investigation, our search for e-cigarette beliefs (and potential relationships of these beliefs to intentions or behaviors) in the literature found little previous work relevant to a youth-focused prevention campaign. More specifically, we searched journals indexed in public health, tobacco control, and communication, in addition to recent conference proceedings (e.g., Society for Research on Nicotine and Tobacco) for studies that explored “electronic cigarettes” or “e-cigarettes” or “vaping.” Citation networks for these studies were then explored to find new studies. We retained studies that explored *beliefs* or *perceptions* that related to e-cigarette use (rather than those that solely explored use patterns or other predictors). In contrast to the 56 studies on traditional cigarettes employed in formative work for *The Real Cost*,⁴ there were only 13 studies that explored the cognitive correlates of e-cigarette use at the time of this investigation. The few studies measuring perceptions about e-cigarettes were largely limited to *relative* harm of e-cigarettes compared to traditional cigarettes rather than the independent, absolute risk of e-cigarette use^{9–12} or surveyed only smokers reporting perceptions about e-cigarettes as smoking cessation devices rather than perceptions of non-users at risk for e-cigarette initiation.^{13–15} Of the small group of studies that did examine youth non-user perceptions, the samples were limited to specific regional areas^{16,17} or focus group research¹⁸, and therefore had limited generalizability for use in a national campaign. Many of these beliefs assessed “harmful”

and “risky,” without specifying actual harms or consequences. Taken together, these early studies of perceptions provided little guidance for specific messages that could be utilized in a campaign or even for communication strategies aimed at e-cigarette prevention, particularly among non-tobacco using adolescents.

A recent benefit to tobacco-control work is the network of Tobacco Centers for Regulatory Science (TCORS) aimed at producing scientific work to inform regulatory decisions related to tobacco. At the time of this study, six TCORS centers were in the process of developing national and state-level representative surveys about e-cigarette perceptions (i.e., Georgia State University, University of California, San Francisco, University of Pennsylvania, University of Texas at Austin, Virginia Commonwealth University, and Yale University). Aggregating the items from these questionnaires comprised a set of 328 items that varied in specificity, absolute risk, and relative risk. For example, some items explored general “risk” or “harm” of e-cigarettes, while others compared that general risk or harm to other products. Some items addressed specific types of e-cigarettes (e.g., tanks vs. mods). Additionally, since the questionnaires were still in the field at the time of our investigation, potential associations between these beliefs and intentions/behaviors were unknown, giving us little guidance for prioritizing the beliefs.

In summary, a prior review of scholarship was sufficient for developing a set of beliefs for *The Real Cost* campaign formative research due to the existence of a consistent, robust body of evidence about the effects of cigarette perceptions on cigarette use. In contrast, a review of scholarship connecting e-cigarette beliefs to e-cigarette use was inadequate because evidence linking perceptions to this novel behavior was limited. Had we relied solely on published work to develop campaign themes, the set of beliefs would be quite small and not relevant for the youth population of interest. Additionally, the dynamic use patterns and perceptions of emerging novel tobacco products might not be captured in scholarship due to the length of the publication process. On the other hand, while the TCORS network provided a large set of potential items, they posed different challenges. Several items were still being tested in the field and in the midst of analysis, and therefore, these items needed supplementary evidence to prioritize certain beliefs over others. Thus, employing a multi-method approach that incorporated elicitation surveys and TM helped address these challenges by prioritizing beliefs uncovered consistently by several methods. These methods also allowed us to detect emerging terms and language associated with these products that the older published studies and survey work might not have included.

Method 2: Elicitation Surveys

Elicitation surveys seek to gather salient attitudinal, normative, and perceived behavioral control beliefs about performing a behavior from the population of interest.¹ Elicitation surveys ask respondents open-ended questions, such as to list the (dis)advantages of their performing a specific behavior. Ideally, these open-ended items provide a list of all possible potential beliefs related to e-cigarettes,¹⁹ which includes those most commonly held (i.e., those most listed by multiple participants) beliefs, but also novel (i.e., emerging perceptions not previously explored) beliefs. Elicitation surveys are particularly useful when respondents are from the target population, so novel beliefs are generated directly from the prospective

audience of the campaign. Their advantage, compared to closed-ended questionnaires, is that they allow respondents free rein to generate ideas. Researchers can identify salient beliefs from elicitation surveys by looking for beliefs that come up more frequently and towards the beginning of the open-ended responses. They are particularly useful, then, for generating a wide range of potential responses. However, they depend on individual differences in articulateness and ability to produce ideas on a topic possibly of low salience; they have a limited claim to providing information about the distribution of responses in a population (exacerbated by the use of a sample of uncertain representativeness.) They share that limitation with focus groups, often used for a similar purpose. However, elicitation surveys (particularly those done on line) are logistically simple to mount (with no need for high cost, synchronous focus group meetings) and are less vulnerable to group dynamics and pressures.

We recruited 176 13–17 year old participants ($M = 15.14$ years; 60.23% non-Hispanic White; 57.43% female) from Toluna, a national non-representative opt-in online survey panel provider, to complete the elicitation survey. 59.7% of the sample had never used an e-cigarette or tobacco cigarette; 22.7% had ever used an e-cigarette, and 23.5% had ever used a tobacco cigarette. Respondents were asked to list in an open-ended format the advantages and disadvantages of using e-cigarettes (behavioral beliefs) and who would (dis)approve of their e-cigarette use (normative beliefs). The most frequently listed advantages were that there are “no advantages,” relaxation/stress relief, and that e-cigarettes are better or healthier relative to tobacco cigarettes. Top disadvantages included addiction, general health risk, and concerns related to breathing/lung health. Interestingly, many groups expected to approve of e-cigarette use were also expected to disapprove (i.e., friends and parents were expected to approve and disapprove of e-cigarette use, respectively).

We noted above that elicitation surveys are subject to concerns if they were meant to provide estimates of distribution of beliefs in a population. However the purpose of this research was to generate a list of potential beliefs of interest worth follow-up systematic research. The criterion for successful elicitation research is theoretical saturation: when eliciting responses from additional respondents no longer produces additional potential beliefs.

However there are other difficulties with this research approach. The open-ended nature of the responses can pose some challenges for researchers. Respondents may use varied descriptions of one belief (e.g., “polluted lungs,” “lung problems,” or “bad for lungs” as lung-related disadvantages of daily e-cigarette use). Slang or spelling/grammar errors may be difficult to interpret, especially when researchers are unable to ask for clarification. For example, one respondent noted that e-cigarettes may be “More hasserdus [*sp*] to your health.”

Elicitation surveys are also subject to constant reevaluation by the research team to assess whether theoretical saturation has been achieved. This requires recognizing when alternative phrasing of beliefs can be merged with an existing belief. Also, emerging behaviors may pose challenges in terms of reaching members of the target population with direct experience with the target behavior. A relatively large number of youth had to be surveyed as

e-cigarette behavior was quite rare, so it was difficult to locate youth that high-rates of use (e.g., daily, as opposed to once or twice a month).

Method 3: Topic Modeling

Another source of e-cigarette perceptions is the mass media which provide depictions of and discussion surrounding e-cigarettes. Gathering beliefs from mass media texts can be useful if plausible and valid interpretations can be made.²⁰ To this end, we used TM, specifically, Latent Dirichlet Allocation (LDA),^{21,22} to identify potential e-cigarette beliefs from a media text database. Put simply, LDA uses modeling assumptions to group patterns of similar terms and phrases from the collection of media texts into topic clusters, which can then be interpreted to develop potential beliefs.²³ The use of statistical modeling reduces the burden of training the number of human coders that would be required to complete this task through traditional content analytic approaches. This method can supplement the more traditional ways of belief generation discussed above and is more sensitive to newly emerging or understudied topics that are broadly discussed in the public arena but not yet researched.

We collected media texts between May 2014 and June 2015 that were centrally devoted to e-cigarette related issues, i.e., contained at least three e-cigarette relevant terms in each text, produced by the Associated Press (AP), the top 50 English-language newspapers published in the United States based on the circulation ranking provided by Lexis-Nexis, and the top 50 most popular websites among 12–24 year olds according to Nielsen's quarterly data. In total, we collected a corpus of more than 4000 texts (AP = 477, newspapers = 1428, websites = 2117). We chose to rely on these lengthier articles instead of shorter social media texts because longer texts are better at providing sufficient density for LDA to return meaningful topic clusters. We then used the Python packages *NLTK*²⁴ and *Ida*²⁵ to conduct the analyses. First, we performed standard preprocessing procedures including tokenization, converting uppercase to lowercase, removing stopwords and punctuation, pruning, and unifying variant terms that refer to the same object (e.g., FDA, F.D.A., Food and Drug Administration). Next in order to generate topic clusters that could be interpreted as arguments and belief themes, we first identified the number of topic clusters for the TM to extract. Generating too few topic clusters may not tap into the nuances of the range of beliefs, while generating too many might yield several semantically redundant clusters, and are prone to pick up noise and irrelevant phrases. We tested models that ranged from 5 to 30 topics in order to explore the range of topics generated. Following standard practice,^{20, 26} three human coders independently reviewed the top 10 terms for each topic generated by these models in order to label each topic cluster. The coders then compared their respective labels and together determined that the 18-topic solution produced the most coherent, efficient, and semantically meaningful topic clusters (Table 1).

For the number of topic clusters specified, TM generates two key pieces of information for each cluster that we used to identify beliefs: 1) a list of the most frequent terms found in each cluster, and 2) a ranking of texts that best represent that topic cluster (e.g., the top document listed would contain the most mentions of the terms in that topic cluster). We then pulled the top 10 most representative texts for each of the 18 topics and looked at the first 1000 characters from each of these 180 documents, since news articles tend to lead with the

most important information.²⁷ Three coders divided the documents and identified (somewhat simultaneously) the potential beliefs or belief themes present in the topic clusters. We collapsed topic clusters that addressed the same theme, added a few new themes that were not captured by initial interpretations based on reading the top-ranked terms only, improved and finalized theme labels, and recorded individual beliefs nested within each theme with some categorized into sub-themes.

While topic modeling provides a valuable and scalable method for exploring the massive public discourse surrounding e-cigarettes, it requires collecting large quantities of texts across media channels. In addition, turning media content into promising targeted belief statements for campaigns is not always intuitive. For example, content such as “Louisiana raises taxes on e-cigarette and vapor products” may not lead to a directly meaningful campaign message. Therefore, gaps remain between using the topic modeling approach to map the dimensions of public discourse and converting topics to readily usable beliefs for campaign planning.

The Benefits of Employing a Multi-Method Approach to Identifying Campaign Themes

The three methods we employed yielded distinct benefits and challenges. However, we argue that it is the complementary nature of the methods that provides the greatest utility for generating potential campaign themes. Through these three methods, we generated a final list of 115 beliefs across 23 belief themes (see Table 2). Specifically, we suggest that a multi-method approach has four key benefits to generating and refining a list of potential message beliefs.

Benefit 1. Developing an exhaustive list of beliefs.

Foremost, data-driven methods such as elicitation surveys or topic modeling provide unique beliefs not found in past literature. For example, the beliefs “I would have fun” (Table 2, belief 25) “it would be easy for me to use them” (Table 2, belief 34), “I would have to spend a lot of money buying the device” (Table 2, belief 42), and “it would be expensive for me” (Table 2, belief 43) were exclusively generated from the elicitation survey. Additionally, public health officials’ concern about e-cigarettes becoming a gateway to illegal drug (e.g., marijuana) use and traditional tobacco use came from the topic modeling but neither appeared in the other methods. Thus, a multi-method approach allows for a more exhaustive pool of potential campaign beliefs.

Benefit 2. Identifying larger message themes, while also exploring nuances.

In addition to providing more beliefs, a multi-method approach allowed us to both identify larger themes and explore nuances within already established themes and beliefs. The elicitation survey identified several dimensions of mood management related to e-cigarettes, for example. Respondents provided advantages such as “relieves stress,” “takes away pain,” “calms you down,” and “it relaxes me”, which provided nuance to the mood management theme that included both relaxation and stress relief. In another instance, while typical measures assessing descriptive norms around e-cigarettes directly ask about the perception

of how prevalent e-cigarette use is around a respondent, topic modeling provided us with various normative statements such as, “The e-cigarette market is growing” and “Vaping is a niche of tech culture.” A multi-method approach allows researchers to explore a diversity of beliefs within larger categories.

Benefit 3. Identifying appropriate language for the target population.

A benefit of the elicitation survey is its open-ended nature, which allows insight into the terminology the target population uses to describe beliefs. Particularly with youth populations who lack the scientific vocabulary and may use colloquial terms, the elicitation surveys provided language for constructing belief statements. For example, the statements “you have to keep buying more juice” (which referred to the cost of needing to buy more cartridges) and “waste money” (which referred to persistent financial cost) were ways to frame the financial cost aspect of e-cigarettes in appropriate language for youth. On the other hand, elicitation surveys and topic modeling may not yield specific belief statements that are measurable, so past literature can help guide development of belief statements from those methods. For example, some belief items were developed by integrating all three sources (1) the language from the elicitation survey “putting chemicals in your body,” (2) specific chemicals that arose from topic modeling such as nickel and propylene glycol, and (3) the question structure from previous survey items in the literature.

Benefit 4. Prioritizing belief themes.

Finally, the three methods provided a certain degree of overlapping beliefs. We prioritized themes found using multiple methods (see Table 3). However, we note that topic modeling, which was based in a virtual census of media content from a wide range of sources during an extended period, and thus might be considered the most representative source of information, provided a particularly useful way to prioritize beliefs. The TM approach gave us topic prevalence rankings indicating which topics are discussed most frequently in the text corpus (e.g., “Research Evidence” and “Comparative Risk” are high-frequency topics in our study; Table 1). On the belief level, for example, both addiction-related statements (e.g., “E-cigarettes could be a portal to nicotine addiction”) and cessation-related statements (e.g., “E-cigarettes help people quit smoking”) were mentioned six times in TM, thereby adding confidence that they are salient beliefs to focus on within broader message themes.

Limitations

Though we have acknowledged several reasons why a multi-method approach is helpful, it is important to detail the limitations of such an approach. A multi-method approach can be costly and time-consuming. It is best employed when there are enough personnel to complete each method concurrently, and ideally, who already possess the training for each method. Training human coders for open-ended elicitation surveys or to interpret topic modeling clusters can take several weeks. We had the limitation of a relatively short time-frame (about three months), therefore our process leveraged resources and expertise that were already at our disposal. For example, we had access to a large-scale database of media texts and did not need to create a new data base. There were costs for mounting the elicitation survey but were lower than the cost required per individual member for

comparable focus groups. The largest costs were associated with staff time to design, implement and analyze the data.

Though we have detailed three specific methods, we argue that using multiple methods of any sort can be useful, particularly when research team members are experts in those methods. Other methods of belief generation may also yield diverse beliefs, such as focus groups and in-depth interviews. Perhaps other media sources for TM would also have yielded more diverse beliefs. For example, social media content may have generated additional relevant beliefs, as youth use such platforms more than they read newspapers. Texts from the media platforms primarily used by the target population are presumably most useful for topic modeling or content analysis to generate relevant beliefs. Our approach was also conducted in a specific context (electronic cigarettes) for a specific population (youth). It may be possible that the information environment for other emerging tobacco-related behaviors could be quite different depending on the population who primarily uses the product. This approach documents existing beliefs in the information environment, however, it does not include novel or alternative beliefs that have not yet diffused through the information environment.

Moreover, these methods lead to an important first step in campaign development by generating a list of testable beliefs. Such a list does not indicate which beliefs would be most persuasive in a campaign nor what the distribution of such beliefs are in the population, and thus requires further evaluation. We undertook a follow-on survey designed to address both of these issues and which served as the source for recommendations to the FDA as it planned its campaign.²⁸ Researchers must weigh these considerations when making decisions in the campaign development process.

Conclusion

In summary, campaigns such as *The Real Cost* that pay close attention to beliefs correlated with intention and behavior may have a greater likelihood for success. Because choosing potential messages to test when developing a campaign can be challenging for novel behaviors such as emerging tobacco product use, a multi-method approach can help generate an appropriate list of potentially promising campaign themes, as well as specific beliefs, that can be tested with members of a target audience. We argue a multi-method approach can provide four utilities when planning a potential health campaign for novel health behaviors: 1) developing an exhaustive and complementary list of beliefs; 2) generating overarching themes and distilling broad themes into more nuanced beliefs; 3) identifying language most relevant to the target population; and 4) prioritizing beliefs for message testing. We believe such an approach is particularly useful for future initiatives aimed at emerging tobacco products that fall under the purview of the U.S. Food and Drug Administration's authority. This process can lead to greater confidence that a full range of themes and beliefs have been considered for eventual campaign messaging.

Conflict of Interest Statement:

Research reported in this publication was supported by the National Cancer Institute (NCI) of the National Institutes of Health (NIH) and FDA Center for Tobacco Products (CTP) under Award Number P50CA179546. The

content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the Food and Drug Administration (FDA). The authors have no other conflicts of interest to disclose.

References

1. Fishbein M, Ajzen I. Predicting and changing behavior: the reasoned action approach. New York, NY: Taylor & Francis; 2010.
2. Rosenstock IM. Historical origins of the health belief model. *Health Educ Monogr.* 1974;2(4):329–335.
3. Brennan E, Gibson LA, Kybert-Momjian A, Liu J, Hornik RC. Promising themes for antismoking campaigns targeting youth and young adults. *Tob Regul Sci.* 2017;3(1):29–46. 10.18001/TRS.3.1.4. [PubMed: 28989949]
4. Brennan E, Momjian A, Jeong M, Naugle D, Parvanta S, Hornik RC. Mass media campaigns to reduce smoking among youth and young adults: documenting potential campaign targets and reviewing the evidence from previous campaigns. CECCR Working Paper Series. Published 6 2012. Philadelphia, PA: Penn's Center of Excellence in Cancer Communication Research, Annenberg School for Communication, University of Pennsylvania.
5. Duke JC, Farrelly MC, Alexander TN, MacMonegle AJ, Zhao X, Allen JA, et al. (2017). Effect of a national tobacco public education campaign on youth's risk perceptions and beliefs about smoking [published online 7 31, 2017]. *Amer J Health Promot.* 10.1177/0890117117720745.
6. Farrelly MC, Duke JC, Nonnemaker J, MacMonegle AJ, Alexander TN, Zhao X, et al. Association between The Real Cost media campaign and smoking initiation among youths—United States, 2014–2016. *MMWR Morb Mortal Wkly Rep.* 2017;66(2):47–50. [PubMed: 28103214]
7. Huang LL, Lazard AJ, Pepper JK, Noar SM, Ranney LM, Goldstein AO. Impact of *The Real Cost* campaign on adolescents' recall, attitudes, and risk perceptions about tobacco use: a national study. *Int J Environ Res Public Health.* 2017;14(1):1–11. doi:10.3390/ijerph14010042.
8. Zhao X, Alexander TN, Hoffman L, Jones C, Delahanty J, Walker M, et al. Youth receptivity to FDA's *The Real Cost* tobacco prevention campaign: evidence from message pretesting. *J Health Commun.* 2016;21(11):1153–1160. doi:10.1080/10810730.2016.1233307. [PubMed: 27736365]
9. Choi K, Forster JL. Beliefs and experimentation with electronic cigarettes: a prospective analysis among young adults. *Am J Prev Med.* 2014;46(2):175–178. [PubMed: 24439352]
10. Grana R, Benowitz N, Glantz SA. E-cigarettes: a scientific review. *Circulation.* 2014;129(19):1972–1986. doi:10.1161/circulationaha.114.007667. [PubMed: 24821826]
11. Saddleson ML, Kozlowski LT, Giovino GA, Hawk LW, Murphy JM, MacLean MG, et al. Risky behaviors, e-cigarette use and susceptibility of use among college students. *Drug Alcohol Depend.* 2015;149:25–30. doi:10.1016/j.drugalcdep.2015.01.001. [PubMed: 25666362]
12. Tan ASL, Bigman CA. E-cigarette awareness and perceived harmfulness: prevalence and associations with smoking-cessation outcomes. *Am J Prev Med.* 2014;47(2):141–149. doi:10.1016/j.amepre.2014.02.011. [PubMed: 24794422]
13. Camenga DR, Cavallo DA, Kong G, Morean M, Connell CM, Simon P, et al. Adolescents' and young adults' perceptions of electronic cigarettes for smoking cessation: a focus group study. *Nicotine Tob Res.* 2015;17(10):1235–1241. doi:10.1093/ntr/ntv020. [PubMed: 25646346]
14. Pepper JK, Brewer NT. Electronic nicotine delivery system (electronic cigarette) awareness, use, reactions and beliefs: a systematic review. *Tob Control.* 2014;23(5):375–384. doi:10.1136/tobaccocontrol-2013-051122. [PubMed: 24259045]
15. Ramo DE, Young-Wolff KC, Prochaska JJ. Prevalence and correlates of electronic-cigarette use in young adults: findings from three studies over five years. *Addict Behav.* 2015;41:142–147. doi:10.1016/j.addbeh.2014.10.019. [PubMed: 25452058]
16. Barrington-Trimis JL, Berhane K, Unger JB, et al. Psychosocial factors associated with adolescent electronic cigarette and cigarette use. *Pediatrics.* 2015;136(2):308–17. doi: 10.1542/peds.2015-0639. [PubMed: 26216326]
17. Roditis ML, Halpern-Felsher B. Adolescents' perceptions of risks and benefits of conventional cigarettes, e-cigarettes, and marijuana: a qualitative analysis. *J Adolesc Health.* 2015;57(2):179–185. doi:10.1016/j.jadohealth.2015.04.002. [PubMed: 26115908]

18. Peters RJ, Meshack A, Lin MT, Hill M, Abughosh S. The social norms and beliefs of teenage male electronic cigarette use. *J Ethn Subst Abuse*. 2013;12(4):300–307. doi:10.1080/15332640.2013.819310. [PubMed: 24215223]
19. Cappella JN, Yzer MC, Fishbein M. Using beliefs about positive and negative consequences as the basis for designing message interventions for lowering risky behavior: toward an integrated approach. In: Romer D, ed. *Reducing adolescent risk: Toward an integrated approach*. Thousand Oaks, CA: Sage; 2003:210–219.
20. DiMaggio P, Nag M, Blei DM. Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. government arts funding. *Poetics*. 2013;41(6):570–606.
21. Blei DM. Probabilistic topic models. *Commun ACM*. 2012;55(4):77–84. doi:10.1145/2133806.2133826.
22. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res*. 2003;3:993–1022.
23. Grimmer J, Stewart BM. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Polit Anal*. 2013;21(3):267–297.
24. Bird S, Klein E, Loper E. *Natural language processing with Python: analyzing text with the natural language toolkit*. Sebastopol, CA: O'Reilly Media, Inc; 2009.
25. Riddell Allen, Hopper Timothy, Grivas Andreas. *Lda: 1.0.4* Zenodo; 2016. doi:10.5281/zenodo.57927.
26. Chang J, Boyd-Graber J, Wang C, Gerrish S, Blei DM. Reading tea leaves: how humans interpret topic models. In: Bengio Y, Schuurmans D, Lafferty J, Williams CKI, Culotta A, eds. *Advances in neural information processing systems*. Cambridge, MA: The MIT Press; 2009:288–296.
27. Errico M, April J, Asch A, Khalfani L, Smith M, Ybarra X. The evolution of the summary news lead. *Media Hist Monogr*. 1997;1(1).
28. Sangalang A, Volinsky A, Yang Q, Liu J, Lee S, Gibson LA, & Hornik R (2016). Identifying Promising Campaign Themes to Prevent Youth Initiation of Electronic Cigarette Use. 1–8. Retrieved from https://repository.upenn.edu/asc_papers/517

Table 1

Topic Clusters, Interpretations, and Highest-Ranking Terms Generated from the 18-Topic Solution Using LDA Topic Modeling

Topic ^a	Interpretation	10 Highest-Ranking Terms for Each Topic
1	<i>Irrelevant</i>	like just people dont time know think way want thats
2	<i>Irrelevant</i>	new said york year general according president news reported states
3	Research Evidence	study ecigarette health said researchers research cancer university tobacco new
4	<i>Irrelevant</i>	said people dont say want going think told just years
5	Comparative Risk	ecigarette health public said use people smoking tobacco harmful smokers
6	Tobacco Industry	tobacco cigarette market said company billion ecigarette sales companies american
7	Targeting Minors	ecigarette nicotine liquid said products devices flavors minors foodanddrugadministration children
8	Cessation Tool	smoking smokers quit nicotine cigarettes tobacco help health cancer habit
9	Teen Vaping Trends	students use school ecigarette teens high percent smoking said survey
10	Local Regulations	state law county public health school officials local department laws
11	Chemicals	ecigarette vapor nicotine chemicals cigarettes devices traditional liquid cigarette smoke
12	FDA Regulations	health foodanddrugadministration public cigarettes companies agency ecigarette regulations rules nicotine
13	Gateway to Drug Use	drug new use vape medical year used high legal users
14	Public Restrictions	smoking ban city public smoke policy tobacco ecigarette use places
15	Novel Product	industry company business big new product make market legal years
16	California's Anti-Vaping Bill	tobacco california age smoking state health products campaign states public
17	E-cigarette Sales	tax state million increase new money states sales percent billion
18	Vaping Prevalence	percent national higher ap survey adults compared states department data

^a: Topics are listed in the order of highest to lowest prevalence ranking (based on percentages of tokens in the corpus). The interpretations of topics based on reading top-ranking terms are tentative. It serves as a general framework or guideline for more solid interpretations based on close reading of the most representative text documents associated with each topic. Some collections of words do not present meaningful and interpretable clusters (e.g., topics 1, 2 and 4), as they may be frequently used and co-occurred words in mass media articles in general. Some topics may be collapsed further based on scrutiny of their representative text documents (e.g., topics 6, 15 and 17).

Table 2

Final List of E-Cigarette Belief Statements Generated Across Three Methods

Theme	Belief Statement ^a
Addiction	
1	...I will not be able to stop if I wanted to.
2	...it will be hard for me to put it down.
3	...I will become addicted.
4	...I will become addicted to nicotine.
Cessation	
5	If I smoked tobacco cigarettes, vaping or using e-cigarettes every day (starting to vape or use e-cigarettes) will help me quit.
6	If I smoked tobacco cigarettes, vaping or using e-cigarettes every day (starting to vape or use e-cigarettes) will work better than other quitting aids in helping me quit.
7	If I smoked tobacco cigarettes, vaping or using e-cigarettes every day (starting to vape or use e-cigarettes) will help me reduce the number of cigarettes I smoke.
Opportunities for Social Interaction	
8	...I will be able to socialize with other people who vape.
9	...it will be a good conversation starter.
10	...I will fit in with my peers.
11	...I will feel like less of an outcast.
12	...I will be accepted by others.
13	...I will be able to share my e-cigarette with friends.
14	...it will help me make friends.
E-Cigarette Specific Risk	
15	...I will be exposed to harmful vapor.
16	...I will have to worry about an e-cigarettes catching fire or exploding.
17	...I will be exposed to dangerous ingredients.
18	...I will worry about liquid chemicals from refills leaking on clothes or furniture.
Cosmetic Effects	
19	...I won't smell like smoke.
20	...I will like the way I looked using them.
21	...I won't stain my fingers or clothes.
22	...I won't have bad breath.
23	...my teeth won't be discolored.
24	...I will be able to keep my weight down.
Enjoyment & Mood	
25	...I will have fun.
26	...I will get a nice buzz.
27	...I will enjoy making vape clouds.
28	...I will enjoy the taste.
29	...it will help me concentrate.

Theme	Belief Statement ^a
30	...I will have something to do with my hands.
31	...I will have something to do when I was bored.
Ease of Use	
32	...it will be convenient to carry them with me.
33	...it will be easy for me to hide them.
34	...it will be easy for me to use them.
Relaxation & Mental Health	
35	...it will relax me.
36	...it will clear my mind.
37	...it will calm my nerves.
38	...it will reduce my stress.
39	...it will be better for reducing stress than medications.
40	...it will help when I am upset or angry about something.
Cost (Financial)	
41	...I will have to spend a lot of money buying refills/juice.
42	...I will have to spend a lot of money buying the device.
43	...it will be expensive for me.
44	...it will be a waste of my money.
Modification	
45	...I will be able to control my level of nicotine exposure.
46	...I will enjoy the fun of being able to DIY (do-it-yourself) my own e-cigarettes.
47	...I will be able to modify all parts of the vaping experience.
48	...I will be able to get an e-cigarette without nicotine.
Health Effects - Short Term	
49	...I will have sinus issues.
50	...I will feel dizzy or have headaches.
51	...I will have a dry, itchy throat.
52	...I will be dehydrated.
53	...I will get a bad cough.
54	...I will have a hard time breathing.
55	...it will decrease my sports performance.
Health Effects - Long Term	
56	...I will get oral (mouth) cancer.
57	...I will get lung cancer.
58	...it will harm my lungs.
59	...I will get sick because it will weaken my immune system.
60	...I will develop sexual and/or fertility problems.
61	...it will change my brain.
Social Perceptions - Pro	

Theme	Belief Statement ^a
62	...I will look cool.
63	...I will look confident.
64	...I will look attractive.
Social Perceptions - Anti	
65	...I will look immature.
66	...I will look stupid.
67	...I will look ridiculous.
Gateway & Polyuse	
68	...I will be more likely to use tobacco cigarettes.
69	...I will be more likely to use marijuana.
70	...I will be more likely to use other drugs.
71	...I will be more likely to use other tobacco products, such as cigarillos and hookah.
Harm to Others	
72	...it will be harmful to my friends' health.
73	...it will be harmful to my family's health.
74	...it will harm others around me.
75	...it will produce secondhand smoke.
76	...it will bother people around me.
77	...it will expose others to chemicals absorbed through the skin.
78	...it will be harmful to the environment.
Policy - Public Restrictions	
79	...I will be able to use them in indoor spaces such as restaurants and theaters.
80	...I will be able to use them in outdoor spaces such as parks.
81	...I will be able to carry them on airplanes.
Policy - Purchase Restrictions	
82	...I will not be allowed to purchase them because I'm too young.
83	...I will have to pay tobacco taxes on them.
84	...I will not be able to purchase them in places near my school or home.
Tobacco Industry	
85	...I will be purchasing products from the same people that make tobacco products.
86	...I will be supporting the tobacco industry.
87	...I will feel manipulated by the tobacco companies.
Technology	
88	...I will feel like I am part of tech culture.
89	...I will be using a futuristic device.
90	...I will feel like I am using a cutting-edge product.
Chemicals	
91	...I will inhale nicotine.
92	...I will inhale poisons.

Theme	Belief Statement ^a
93	...I will be exposed to toxic chemicals.
94	...I will be exposed to toxic metals such as chromium, nickel, and lead.
95	...I will be exposed to propylene glycol, which can lead to skin irritation.
96	...I will be exposed to formaldehyde, which can lead to eye, nose, and throat irritation.
97	...I will be exposed to diacetyl, which can lead to “popcorn lung” (lung scarring).
98	...I will be exposed to aerosol, which may contain harmful particles.
99	...I will be exposed to hormones.
100	...I will be exposed to charcoal.
101	...I will be exposed to tar.
Flavors	
102	...I will be able to use a variety of flavors I like.
103	...I will enjoy trying different e-cigarette products and flavors with friends.
104	...the flavor additives will not harm me.
Comparison to Cigarettes	
105	...it will not bother people around me as much as tobacco cigarettes do.
106	...I will avoid chemicals found in tobacco cigarettes.
107	...it will be cleaner than smoking tobacco cigarettes.
108	...I will be able to get them more easily than tobacco cigarettes.
109	...they will be cheaper than smoking tobacco cigarettes.
110	...it will be less harmful than smoking tobacco cigarettes.
111	...it will taste better than smoking tobacco cigarettes.
112	...it will be less harmful to others than smoking tobacco cigarettes.
113	...I will be able to use e-cigarettes where tobacco cigarette smoking is not allowed.
114	...I will not be exposed to the tar found in tobacco cigarettes.
115	...it will be less addictive than tobacco cigarettes.

^aIndividuals hold specific beliefs about a behavior,¹ though sets of beliefs can be collapsed into a larger campaign theme. Different messages can address specific beliefs, but several messages can address the larger theme. For example, the effects of cigarette smoking on “headaches,” “lung disease,” and “cancer” could all be collapsed into the larger theme of “health effects.” Our work explores identifying larger campaign themes (which could represent a whole campaign strategy) and specific beliefs (which could underpin distinct messages). Therefore, we refer to theme-level and belief-level separately.

^bThere were two stems that preceded each belief question: “If I vape or use e-cigarettes every day” or “If I start vaping or using e-cigarettes,” followed by the benefit or negative consequence (e.g., “...it will harm others around me”). Each participant was randomly assigned to one of those stems for all of the belief statements. In the analyses the responses from both stems were collapsed, as preliminary analyses indicated that the two stem conditions produced results that were largely similar to one another, and by combining the data, we were able to increase the sample size and hence the stability of our results.

Table 3

Presence of Potential Campaign Themes Across Methods

Theme		# of Items	Expert Sources		Data-Driven	
			Literature/Past Studies	TCORS/FDA	Elicitation	Topic Modeling
1	Addiction	4	X	X	X	X
2	Cessation	3	X	X	X	X
3	Opportunities for Social Interaction	7	X	X	X	X
4	E-Cigarette Specific Risk	4	X		X	X
5	Cosmetic Effects	6	X	X	X	
6	Enjoyment & Mood	7	X	X	X	
7	Ease of Use	3	X	X	X	
8	Relaxation & Mental Health	7	X	X	X	X
9	Cost	3			X	
10	Modification	4	X	X		
11	Health Effects - Short Term	7	X	X	X	X
12	Health Effects - Long Term	7	X	X	X	X
13A	Social Perceptions - Pro	3	X	X	X	X
13B	Social Perceptions - Anti	3	X	X	X	
14	Gateway & Polyuse	4	X			X
15	Harm to Others	7	X	X	X	X
16	Policy - Public Restrictions	3		X		X
17	Policy - Purchase Restrictions	3	X	X		X
18	Tobacco Industry	3		X		X
19	Technology	3		X		X
20	Chemicals	11		X		X
21	Flavors	3	X	X	X	X
22	Comparison to Cigarettes	10	X	X	X	X

Note. The belief generation yielded all perceptions that were both pro- or anti-use. In future testing, all belief wording should be altered to yield the same direction of valence within category.