



Published in final edited form as:

*Tob Prev Cessat.* 2018 February ; 4: . doi:10.18332/tpc/84866.

## Validation of mobile phone text messages for nicotine and tobacco risk communication among college students: A content analysis

Georges E. Khalil<sup>1</sup>, Karen S. Calabro<sup>1</sup>, Brittani Crook<sup>2</sup>, Tamara C. Machado<sup>1</sup>, Cheryl L. Perry<sup>2</sup>, and Alexander V. Prokhorov<sup>1</sup>

<sup>1</sup>The University of Texas MD Anderson Cancer Center, Texas, United States

<sup>2</sup>The University of Texas Health Science Center, School of Public Health, Texas, United States

### Abstract

**INTRODUCTION**—In the United States, young adults have the highest prevalence of tobacco use. The dissemination of mobile phone text messages is a growing strategy for tobacco risk communication among young adults. However, little has been done concerning the design and validation of such text messages. The Texas Tobacco Center of Regulatory Science (Texas-TCORS) has developed a library of messages based on framing (gain- or loss-framed), depth (simple or complex) and appeal (emotional or rational). This study validated the library based on depth and appeal, identified text messages that may need improvement, and explored new themes.

**METHODS**—The library formed the study sample (N=976 messages). The Linguistic Inquiry and Word Count (LIWC) software of 2015 was used to code for word count, word length and frequency of emotional and cognitive words. Analyses of variance, logistic regression and scatter plots were conducted for validation.

**RESULTS**—In all, 874 messages agreed with LIWC-coding. Several messages did not agree with LIWC. Ten messages designed to be complex indicated simplicity, while 51 messages designed to be rational exhibited no cognitive words. New relevant themes were identified, such as health (e.g. ‘diagnosis’, ‘cancer’), death (e.g. ‘dead’, ‘lethal’) and social connotations (e.g. ‘parents’, ‘friends’).

**CONCLUSIONS**—Nicotine and tobacco researchers can safely use, for young adults, messages from the Texas-TCORS library to convey information in the intended style. Future work may expand upon the new themes. Findings will be utilized to develop new campaigns, so that risks of nicotine and tobacco products can be widely disseminated.

---

This is an Open Access article distributed under the terms of the Creative Commons Attribution NonCommercial 4.0 International License. (<http://creativecommons.org/licenses/by-nc/4.0>)

CORRESPONDENCE TO: Georges Elias Khalil. The University of Texas MD Anderson Cancer Center, 1155 Pressler Street Unit 1330, 77030 Houston, United States. [georgesdoc@gmail.com](mailto:georgesdoc@gmail.com).

#### CONFLICTS OF INTEREST

Authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none was reported.

#### PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed

## Keywords

tobacco; nicotine; validation; risk communication; text message; message framing

---

## INTRODUCTION

Current research in the United States indicates that young adults (18–24 years) have the highest prevalence of using any type of tobacco product (37.6%)<sup>1,2</sup>. Young adults are at risk of using conventional tobacco products (e.g. cigarettes), new and emerging tobacco products (e.g. e-cigarettes), and multiple products at the same time<sup>3–5</sup>. In particular, young adults in community college represent an underserved population that is more likely to use tobacco than young adults attending 4-year level universities<sup>6–8</sup>. Tobacco use among young adults can be attributed to the relatively lower knowledge and perception of the risks of tobacco, when compared to other adult age groups<sup>9–16</sup>. As a result, there is a need to engage in campaigns that communicate to young adults about tobacco risks<sup>17,18</sup>.

The dissemination of text messages via mobile phones is a growing strategy for tobacco-risk communication<sup>9,19</sup>. Considering that 96% of American young adults own mobile phones capable of receiving texts, text messaging is a particularly appropriate method for transmitting information to this population<sup>20,21</sup>. While this strategy has been successful<sup>22–24</sup>, little has been done to design validated text messages for tobacco-risk communication. Furthermore, a validated library of text messages has yet to be reported in the literature.

Researchers have constructed text messages for risk communication based on three main structures: framing (gain-framed or loss-framed messages), depth (simple or complex messages), and appeal (emotional or rational messages)<sup>25–28</sup>. Compared to depth or appeal, developing text messages based on framing has been relatively easy due to the well-established conceptualization of gain- and loss-framing<sup>29,30</sup>. In short, gain-framed messages emphasize the benefits of quitting or avoiding substance use, while loss-framed messages emphasize the costs of use. Such a conceptualization of framing has been consistently applied by researchers in health promotion and disease prevention<sup>31</sup>.

Researchers face challenges when developing messages based on depth and appeal. The diversity of message features that depict depth and appeal makes it difficult to effectively construct a text message. In the context of depth, message complexity has been defined in many ways, based on message structure, content, or both<sup>32</sup>. While structure may involve complex grammatical applications, content can include longer words<sup>33–35</sup>. Research suggests that message complexity can have an effect on tobacco-risk communication, and the success of such messages depends on an individual's level of need for cognition<sup>36</sup>. Individuals with low need for cognition are more likely to express intention to quit when exposed to simple messages, whereas individuals with high need for cognition are more likely to be persuaded by complex messages<sup>36</sup>. In terms of appeal, research on text message development has conceptualized emotional messages based on linguistic cues and paralinguistic cues. Linguistic cues include the use of emotional words (e.g. 'happy', 'angry')<sup>37</sup> and linguistic markers (i.e. expressing emotion without emotional words, e.g. 'I

want to thank you so much')<sup>37–39</sup>. Paralinguistic cues in text messages express nonverbal cues that are normally communicated physically. There are five types of paralinguistic cues: *vocal spelling* (mimicking a specific vocal inflection, e.g. 'weeeell', 'soooo'), *lexical surrogates* (textual representations of vocal sounds that are not words, e.g. 'uh-huh', 'haha'), *spatial arrays* (pictographs constructed from punctuation and letters, e.g. :-) for happy face), *manipulation of grammatical markers* (alterations of the presentation of words, e.g. all capital letters, strings of periods or commas), and *minus features* (deliberate or inadvertent neglect of conventional formatting elements, e.g. lack of capitalization or paragraphing)<sup>40</sup>. Previous studies have shown that emotionally evocative messages using fear appeal or humor can be successful for tobacco-risk communication, as they facilitate recall of message content, increase tobacco-related knowledge and motivating users to quit the use of tobacco<sup>27,41,42</sup>.

Beyond manual analysis of text, communication researchers have worked to develop software programs that can automatically analyze the content of text. Such programs categorize message content based on themes that have been previously identified through extensive traditional manual coding. In particular, we focused on a method, the Linguistic Inquiry and Word Count (LIWC), computerized text-analysis software, that counts the frequency of words and word stems to study the emotional, cognitive, structural and process components in written text or speech<sup>43–45</sup>. Considering the complexity of message design with respect to depth and appeal, the LIWC procedure can automatically and quantitatively identify text message features that allow the differentiation between message depth and appeal. This study is the first to apply the LIWC procedure in order to validate text messages in the context of tobacco research.

At the Texas Tobacco Center of Regulatory Science (Texas-TCORS), our researchers have developed a library of mobile phone text messages categorized on framing, depth and appeal. The objective of the text messages is to communicate the risks of tobacco use to young adults, both users and nonusers. While the Texas-TCORS library of text messages has been developed through extensive formative research, message categorization has yet to be objectively validated. Predictive validity of the library is crucial as it will ensure that the messages are correctly designed based on their category, and it will allow researchers to conduct randomized trials based on the message categories with confidence. Results from trials can guide potentially more effective communication campaigns. In addition, as we examine message content, we may be able to extract new themes emerging from the messages. Using the LIWC procedure, this paper aims to: 1) validate the library of gain- and loss-framed text messages based on appeal and depth, 2) identify messages that may be improved, and 3) explore additional categories of message design in the Texas-TCORS text message library.

## METHODS

### Text message development

From January 2014 to August 2015, the Texas-TCORS researchers developed 976 messages, taking into account previous scientific literature, trends in social media related to tobacco product use and trending terminology. Collectively, the research team has extensive

experience in tobacco-risk communication, public health, psychology and creative writing. Message development also involved focus groups conducted with young adults, and external experts in health communication, tobacco control and public health. Overall, the focus group discussions indicated that the messages were perceived as interesting and appropriate. Feedback on the messages was incorporated into message revisions, to reach a final version of the messages<sup>46</sup>.

The library included 976 text messages, developed according to framing, depth and appeal. A permutation of the three structures was implemented to have eight categories of text messages (2×2×2): gain-framed/simple/emotional, loss-framed/simple/emotional, gain-framed/complex/emotional, loss-framed/complex/emotional, gain-framed/simple/rational, loss-framed/simple/rational, gain-framed/complex/rational, and loss-framed/complex/rational (122 messages per category).

Messages describing conventional products included information about combustible cigarettes, cigars, cigarillos, smokeless tobacco and pipes. Messages about new and emerging products included information about e-cigarettes, vapes, electronic liquids and hookah (waterpipes).

### Study sample

The library comprised the study sample (N=976 messages). The unit of analysis was the individual text message, defined as what one individual may receive via mobile phone. This included the total content of the post, regardless of length of a given text message (from a single word to multiple paragraphs). Half of the messages were gain-framed and the others were loss-framed. The number of characters per message varied between 52 and 172 characters. Examples of text messages as categorized initially by writers are presented in Table 1.

### Content analysis

The messages were coded using LIWC, which is a valid and reliable method for content analysis of text<sup>47</sup>. The LIWC, software version 2015, codes for 103 variables using a full dictionary of words<sup>48</sup>. LIWC calculates ratios or percentages of words that tap on a specific variable (i.e. the number of category words in a single message, divided by the total number of words in that message). These ratios, as opposed to word counts, were used to account for the difference between messages with respect to the amount of content<sup>45</sup>.

### Measures

Using LIWC, we measured message characteristics indicative of depth and appeal, as supported by previous literature<sup>3,37,49–54</sup>. Message depth was measured through: word length (i.e. frequency of words with 6 letters or more) and word count (i.e. number of words per message)<sup>52,53</sup>. Message appeal was measured using characteristics that distinguish between emotional and rational messages. To capture emotionality, we measured affect (i.e. frequency of words expressing overall emotion, such as ‘cheerful’, ‘hopeful’, and ‘humor’) and subcategories of affect including negative emotions (e.g. ‘hurt’, ‘mad’, and ‘risk’), positive emotions (e.g. ‘happy’, ‘cheerful’, and ‘thankful’), anger (e.g. ‘rage’, ‘anger’, and

‘aggressive’), anxiety/fear (e.g. ‘anxious’, ‘avoid’, and ‘afraid’), and sadness (‘sad’, ‘alone’, and ‘cry’)<sup>37</sup>. To capture rationality, we measured two variables: cognitive processing (i.e. frequency of words depicting cognitive processing such as ‘think’, ‘decide’, and ‘perhaps’) and quantification (frequency of words depicting amounts such as numbers, ‘much’, ‘many’, and ‘few’)<sup>54</sup>.

Finally, we coded for nine new themes using LIWC, based on common themes tackled when communicating tobacco risks<sup>3,49–51</sup>. In particular, we identified terms related to: health (e.g. ‘diagnosis’, ‘healthy’, and ‘cancer’), death (e.g. ‘death’ and ‘lethal’), social (e.g. ‘parents’, ‘kids’, and ‘friends’), leisure (e.g. ‘bar’ and ‘restaurant’), religion (e.g. ‘God’, ‘pray’, and ‘blessing’), body (e.g. ‘lungs’, ‘skin’, and ‘heart’), work and marketing (e.g. ‘company’ and ‘job’), money (e.g. ‘cost’ and ‘buy’), and sexuality (e.g. ‘pregnant’ and ‘erection’).

## Data analysis

We analyzed the text messages separately for depth and appeal. In order to validate the messages based on depth, we first conducted one-way analysis of variance (ANOVA) to determine if message categorization based on depth is related to higher scores on word count and word length. We compared the messages designed by writers to be complex with messages designed to be simple with respect to word count and word length. Then, we conducted multiple logistic regression analysis predicting message design as simple, with word count and word length as the main independent variables, controlling for appeal, framing, affect, cognitive processing and type of nicotine/tobacco product per message.

In order to validate the messages based on appeal, we conducted one-way ANOVA to determine if message categorization by writers is related to higher scores on emotional and rational variables identified by LIWC. We compared the messages designed by writers to be emotional to the messages designed to be rational with respect to affect, stress, anxiety/fear, anger, cognitive processing and quantity. We conducted multiple logistic regression analysis predicting message design as emotional, with affect and cognitive processing as the main independent variables, controlling for depth, framing, word count, word length and type of nicotine/tobacco product per message. With ANOVA, Bonferroni adjustment corrected for alpha over-repeated comparisons and guarded against type 1 error<sup>55–57</sup>.

To identify messages showing disagreement between writers’ design and objective coding of depth, we conducted a scatter plot of word count versus word length, stratifying between messages designed to be simple and messages designed to be complex. Messages designed to be simple with word length and word count above the medians were identified as messages that disagree with objective coding through LIWC. Also, messages designed to be complex with word length and word count below the medians were identified as messages that disagree with objective coding through LIWC.

Similarly, to identify the messages that indicate disagreement between categorization by writers and objective coding of appeal, we presented a scatter plot of affect versus cognitive processing, stratifying between messages designed to be emotional and messages designed to be rational. Messages designed to be emotional with cognitive processing higher than the median and affect lower than the median were identified as messages that disagree with

objective coding by LIWC. Also, messages designed to be rational with cognitive processing lower than the median and affect higher than the median were identified as messages that do not match objective coding.

## RESULTS

We conducted descriptive statistics for the variables of interest after the LIWC procedure, for both loss-framed and gain-framed messages (Table 2). Word count, frequency of affect, positive emotions and anxiety/fear were significantly different between gain-framed and loss-framed messages.

### Validity of message categorization based on depth

Message depth validity indicated that the messages designed to be complex had a significantly higher number of words per message ( $F[1, 974]=72.80, p<0.001, \eta^2=0.07$ ) and frequency of words over six letters ( $F[1, 974]=562.25, p<0.001, \eta^2=0.37$ ). Messages designed to be complex were more likely to present longer words and higher number of words than messages designed to be simple. Supporting such results, logistic regression analysis indicated that the higher the word count and word length in messages the more likely they are designed as complex, controlling for message framing, message appeal, affect, cognitive processing and type of nicotine/tobacco product mentioned in the message (Table 3).

### Validity of message categorization based on appeal

Messages designed to be emotional were more likely to exhibit words of affect, positive emotions, negative emotions, anxiety and anger, than messages designed to be rational (Table 4). There was no significant difference between rational and emotional messages, with respect to words indicating sadness. Messages designed to be rational were more likely to present words using cognitive processing and quantity than messages designed to be emotional (Table 4).

Logistic regression analysis indicated that higher frequency of affect and lower frequency of cognitive processing words were related to more emotional messages, controlling for message framing, message depth, word count, word length and type of nicotine/tobacco product mentioned in the message (Table 5). Among all emotion-related variables, affect presented the highest variance (21%) in predicting message categorization by writers. Between the two cognitive-related variables, cognitive processing presented higher variance (3%) in predicting message categorization by writers.

### Identification of messages needing improvement

Overall, scatter plot results indicated consistency between categorization by writers and LIWC-coding of depth and appeal features (Fig. 1). Messages scattered above the medians in word length (median=27) and word count (median=23) (Fig. 1a, Quadrant 2) were all designed to be complex ( $n=258$ ; 52.87% of all complex messages). On the other hand, 2.0% of complex messages ( $n=10$ ) and 43.03% of simple messages ( $n=210$ ) were scattered below the medians in word length and word count (Fig. 1a, Quadrant 3). In other words, 95.45% of



text messages scattered in this quadrant were designed to be simple. The 10 complex messages scattered below the medians deserve further improvement. However, such messages are very close to the medians (Fig. 1a). One example is ‘Cigarettes may be appealing to some youth, however young people who do not smoke are more likely to avoid impaired cognition’ (Word count=21 words, Word length=23.81% of words over six letters). There was no message designed to be simple that scored above the medians in word count and word length (Quadrant 1). However, Figure 1a indicates 18.65% of simple messages scored above the median in word length alone (Quadrants 1 and 2). One example is ‘Menthol-release cigs [cigarettes] let ppl [people] add menthol to regulars, by pinching a bead. Smoking them makes ppl targets of the tobacco industry’, categorized by writers as loss-framed/simple/rational (Word count=22 words, Word length=40.91% of words over six letters). Also, 19.88% of complex messages (n=97) scored below the median in word length, despite high scores in word count. One example is ‘Oh dear! Lung cancer is always a threat to be concerned about even with “light” cigarettes. All cigarettes prey on the body’s organs and leave them to rot :(’, categorized by writers as loss-framed/complex/emotional (Word count=29 words, Word length=10.34% of words that over six letters).

In the context of message appeal (Fig. 1b), several messages agreed with LIWC coding. According to the findings, 41.60% of rational messages (n=203) and 28.69% of emotional messages (n=70) were scattered below the median in affect (median=9.55), and above the median in cognitive processing (median=9.50) (Fig. 1b, Quadrant 4). In other words, 74.36% of text messages scattered in this quadrant are designed by writers as rational. Even though the 70 emotional messages were in quadrant 4, they all presented some level of affect. None of the messages designed to be emotional scored zero on affect (M=13.94, SD=6.38, range score in affect 3.45-36.36).

On the other hand, 41.60% of emotional messages (n=203) and 28.69% of rational messages (n=70) were scattered above the median in affect, and below the median in cognitive processing (Fig. 1b, Quadrant 1). In other words, 74.36% of text messages scattered in this quadrant are designed by writers as emotional. The 70 rational messages in quadrant 1 may need further improvement, considering that 51 of these messages scored zero on cognitive processing. Also, 89% of the 51 messages exhibited words of affect such as ‘help’, ‘avoid’ and ‘risk’. Two messages are ‘Avoiding hookah prevents exposure to high amounts of benzene (a carcinogen), which decreases the risk of developing acute non-lymphocytic leukemia (ANLL)’ and ‘By avoiding cigarettes, young people reduce their risk of developing Buerger’s disease and maintain healthy circulation to extremities’ (Fig. 1b, Quadrant 1).

The scatter plot also indicated that 377 (77.25%) of the messages that were designed to be rational were coded by LIWC with some level of affect. These messages used words such as ‘bad’, ‘dangerous’ and ‘serious’. Examples of messages are ‘MYTH: Menthols aren’t as bad as regular cigs. TRUTH: Smoking either is dangerous’, categorized by writers as loss-framed/simple/rational (frequency of affect words=23.08%, frequency of cognitive processing words=23.08%) and ‘Drinking e-cig juice can cause serious nicotine poisoning. Toddlers who live around e-cig users are at risk of poisoning’, categorized by writers as loss-

framed/simple/rational (frequency of affect words=21.05%, frequency of cognitive processing words=5.26%).

### **New categories of messages in the Texas-TCORS Library**

We coded for nine themes for the Texas-TCORS Library, describing common topics for risk communication (Table 5). The most common theme is the description of social situations and social connections, with words such as ‘parents’, ‘kids’ and ‘friends’. This theme is coded separately from leisure, which involves terms such as ‘bars’ and ‘hanging out’. The least common theme is the use of religious terms such as ‘blessing’, ‘pray’ and ‘demon’ (Table 6).

## **DISCUSSION**

This is the first report on the design and validation of a large-scale text message library for tobacco-risk communication. Overall, the results indicate that: 1) the majority of the text messages of the Texas-TCORS library are valid messages to communicate risk and used a wide range of depth and appeal, 2) several identified text messages may benefit from improvement, and 3) the library contains themes beyond framing, depth or appeal.

Based on these findings, tobacco-control advocates who intend to disseminate mobile phone text messages for risk communication can safely apply several messages from the library to young adults to convey information in the intended style, whether emotional, rational, simple or complex. In all, 874 (89.55%) text messages were found to be appropriately associated with depth and appeal factors. Thus, for the overwhelming majority of the text messages, predictive validity indicated that message development based on depth and appeal agrees with LIWC-coding. The higher the frequency of long words and the higher the number of words per message the more likely the message is complex. Also, the higher the frequency of affect words and the lower the frequency of cognitive words the more likely a message is emotional. The results are further supported by visually inspecting scatter plots. The medians for word count and word length corresponded to approximately half complex messages and 40% simple messages. Similarly, the medians for affect and cognitive processing corresponded to approximately 40% emotional messages and 40% rational messages.

However, several messages presented disagreement between LIWC-coding and writers’ design and may benefit from additional modifications in order to be used during tobacco-control campaigns based on their intended style. In the context of message depth, 10 messages designed to be complex indicated simplicity based on word length and word count. Disagreement in message depth was mainly due to the use of long words. About 18% of messages, designed to be simple, scored above the median in word length. The use of shorter words may assist in making such messages simpler. Several messages may need rectification with respect to appeal. In particular, 51 messages designed to be rational exhibited no cognitive processing words. Cognitive processing words (e.g. ‘thus’, ‘if’, ‘because’, ‘perhaps’) can allow the messages to present logic in a chain of thoughts, and as a result strengthen their rational style. These revised messages can then be re-examined within the LIWC coding framework to ensure they are consistent with intended themes. In addition,



gain-framed and loss-framed messages differed with respect to emotional appeal and word count. As a result, researchers may need to reexamine framing and balancing the two categories with respect to appeal and complexity.

While the messages were purposely designed based on framing, depth and appeal, new themes were identified during LIWC coding. In addition to themes such as health and death, typical themes relevant for tobacco-risk communication were found, including social connections and leisure, work and marketing, money, sexual connotations, and religion. Ultimately, tobacco-control professionals can select text messages based on the themes that can answer to their populations needs. Further categorization may give mobile phone campaigns a higher level of tailoring, based on college student beliefs, risk perceptions and misconceptions about the effects of tobacco.

Our application of LIWC for content analysis has some assumptions and associated limitations. First, compared with traditional manual coding, LIWC content analysis provides an objective quantification of message depth and appeal. However, this crucially assumes that: 1) the word count and word length are accurate proxies for message complexity, and 2) the frequencies of affect and cognitive processing words are appropriate proxies for message appeal. As a result, some other characteristics of message depth and appeal are not considered, such as reading competency and paralinguistic cues of emotion expression. Nevertheless, the coding features used by LIWC have been supported by previous literature<sup>37,52–54</sup>, and were able to predict message categories as designed by writers. Second, being an automated procedure, LIWC-coding does involve limitations. While LIWC captures frequency of words for a specific category, it does not capture insinuations made through a group of words in a message. This could be addressed by researchers through manual coding. Third, this study is limited to English-speaking young adults attending community colleges. In the future, it may be beneficial to conduct an intensive work of cross validation with a wider audience, including different cultures of young adults and in the context of other languages.

Several implications can be inferred from this study. First, the findings suggest that tobacco control professionals who aim to engage in risk communication need to develop health messages with careful consideration of content and the application of several design approaches. In addition to social and behavioral formative research (e.g. focus group discussions, in-depth interviews, application of health theories, and literature review), the use of an objective content analysis method, such as LIWC, can benefit message development at several levels. In particular, by being an objective and evidence-based method of text coding, the LIWC procedure can create a foundation of data about the text messages among tobacco-control professionals as they make decisions about message content and features. Also, future work may consider expanding upon the themes identified in the current study. For instance, developing text messages that include social connotations and leisure may be useful in order to tap on the role of social influence and peer-pressure in tobacco use among youths. As a next step, the Texas-TCORS library will be tested as part of a randomized controlled trial to determine the most effective text messages with young adults, in terms of increasing risk perceptions. Findings will be utilized as the basis for other

campaigns with this age group and other populations, so that the risks of conventional and emerging tobacco products can be widely disseminated.

## CONCLUSIONS

Tobacco-risk researchers can safely use messages from the Texas-TCORS library with young adults to convey information in the intended style. While some messages may benefit from additional modifications, most revealed agreement between LIWC and human categorization. In addition, several new themes are identified from the message library using LIWC, including social connections and leisure, work and marketing, money, sexual connotations, and religion. Future work may expand upon the new themes. Findings will be utilized to develop new campaigns, so that risks of tobacco products can be widely disseminated.

## Acknowledgments

We acknowledge the work of the scientific committee of the Texas-TCORS and the team of text message developers, including but not limited to Damon Vidrine, Keryn E Pasch, Tamara Machado, Sophia Russell, Katarzyna Czerniak and Gabrielle Botello.

### FUNDING

This work was supported by grant number [1 P50 CA180906] from the National Cancer Institute and the FDA Center for Tobacco Products (CTP). 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.

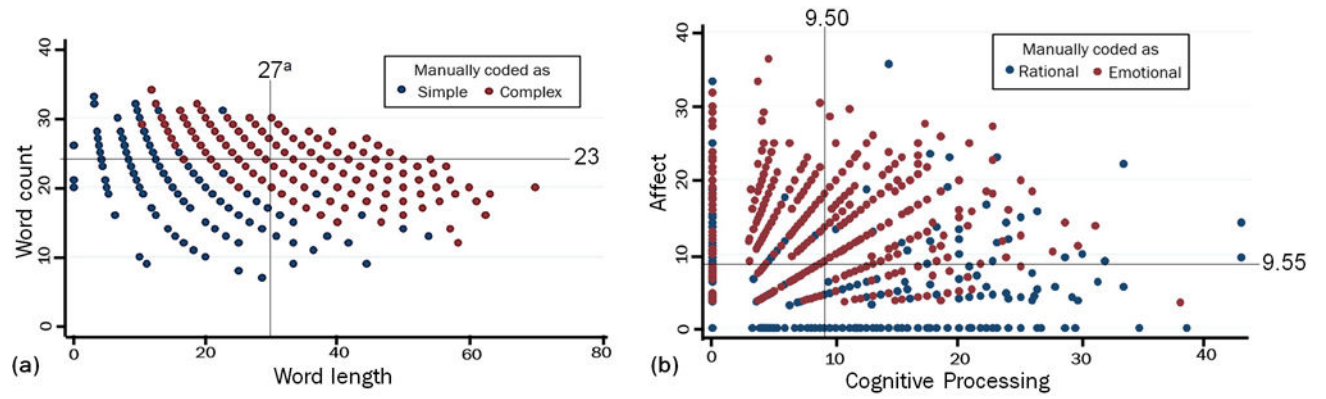
## References

1. Rudolph MD, Miranda-Domínguez O, Cohen AO, Breiner K, Steinberg L, Bonnie RJ, et al. At risk of being risky: The relationship between ‘brain age’ under emotional states and risk preference. *Dev Cogn Neuros*. 2017; 24:93–106. DOI: 10.1016/j.dcn.2017.01.010
2. Kasza KA, Ambrose BK, Conway KP, Borek N, Taylor K, Goniewicz ML, et al. Tobacco-Product Use by Adults and Youths in the United States in 2013 and 2014. *N Engl J Med*. 2017; 376(4):342–353. DOI: 10.1056/NEJMsa1607538 [PubMed: 28121512]
3. US Department of Health and Human Services. The health consequences of smoking—50 years of progress: a report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014. p. 17
4. Rath JM, Villanti AC, Abrams DB, Vallone DM. Patterns of tobacco use and dual use in US young adults: the missing link between youth prevention and adult cessation. *J Environ Public Health*. 2012; 2012:1–9. DOI: 10.1155/2012/679134
5. Soneji S, Sargent J, Tanski S. Multiple tobacco product use among US adolescents and young adults. *Tob control*. 2016; 25(2):174–180. DOI: 10.1136/tobaccocontrol-2014-051638 [PubMed: 25361744]
6. Loukas A, Murphy JL, Gottlieb NH. Cigarette smoking and cessation among trade or technical school students in Texas. *J Am Coll Health*. 2008; 56(4):401–7. DOI: 10.3200/JACH.56.4.401-408 [PubMed: 18316284]
7. O’Brien F, Simons-Morton B, Chaurasia A, Luk J, Haynie D, Liu D. Post-High School Changes in Tobacco and Cannabis Use in the United States. *Subst Use Misuse*. 2017; 53(1):1–10. DOI: 10.1080/10826084.2017.1322983 [PubMed: 28742416]
8. American Association of Community Colleges. Financially challenged: How Often Students Ran Out of Money Over The Past Year. Data Points: Financially challenged. Jan 8, 2018 <http://www.webcitation.org/6wJwkQR5e>. March 16, 2017

9. Berg CJ, Stratton E, Schauer GL, et al. Perceived harm, addictiveness, and social acceptability of tobacco products and marijuana among young adults: marijuana, hookah, and electronic cigarettes win. *Subst Use Misuse*. 2015; 50(1):79–89. DOI: 10.3109/10826084.2014.958857 [PubMed: 25268294]
10. Berry C, Burton S, Howlett E. Are Cigarette Smokers', E-Cigarette Users', and Dual-Users' Health Risk Beliefs and Responses to Advertising Influenced by Addiction Warnings and Product Type? *Nicotine Tob Res*. 2017; 19(10):1185–1191. DOI: 10.1093/ntr/ntx075 [PubMed: 28379568]
11. Wagoner KG, Cornacchione J, Wiseman KD, Teal R, Moracco KE, Sutfin EL. E-cigarettes, hookah pens and vapes: adolescent and young adult perceptions of electronic nicotine delivery systems. *Nicotine Tob Res*. 2016; 18(10):2006–2012. DOI: 10.1093/ntr/ntw095 [PubMed: 27029821]
12. Rodu B, Plurphanswat N, Hughes JR, Fagerström K. Associations of Proposed Relative-Risk Warning Labels for Snus With Perceptions and Behavioral Intentions Among Tobacco Users and Nonusers. *Nicotine Tob Res*. 2016; 8(5):809–816. DOI: 10.1093/ntr/ntv168
13. McDonald EA, Ling PM. One of several 'toys' for smoking: young adult experiences with electronic cigarettes in New York City. *Tob Control*. 2015; 24(6):588–593. DOI: 10.1136/tobaccocontrol-2014-051743 [PubMed: 25564287]
14. O'Brien EK, Nguyen AB, Persoskie A, Hoffman ACUS. Adults' addiction and harm beliefs about nicotine and low nicotine cigarettes. *Prev Med*. 2017; 96:94–100. DOI: 10.1016/j.ypmed.2016.12.048 [PubMed: 28034733]
15. Tan AS, Mello S, Sanders-Jackson A, Bigman CA. Knowledge about Chemicals in e-Cigarette Secondhand Vapor and Perceived Harms of Exposure among a National Sample of US Adults. *Risk Analysis*. 2017; 37(6):1170–1180. DOI: 10.1111/risa.12676 [PubMed: 27595498]
16. Gowin M, Cheney MK, Wann TF. Knowledge and Beliefs About E-Cigarettes in Straight-to-Work Young Adults. *Nicotine Tob Res*. 2017; 19(2):208–214. DOI: 10.1093/ntr/ntw195 [PubMed: 27613919]
17. US Food and Drug Administration. Research Priorities. Tobacco Products. Jan 1, 2018 <http://www.webcitation.org/6wJyzWzzm>. January 5, 2018
18. Ashley DL, Backinger CL, van Bommel DM, Neveleff DJ. Tobacco regulatory science: research to inform regulatory action at the Food and Drug Administration's Center for Tobacco Products. *Nicotine Tob Res*. 2014; 16(8):1045–1049. DOI: 10.1093/ntr/ntu038 [PubMed: 24638850]
19. Leischow SJ, Zeller M, Backinger CL. Research priorities and infrastructure needs of the Family Smoking Prevention and Tobacco Control Act: science to inform FDA policy. *Nicotine Tob Res*. 2012; 14(1):1–6. DOI: 10.1093/ntr/ntn057 [PubMed: 22184581]
20. Rainie, L., Perrin, A. 10 facts about smartphones as the iPhone turns. Aug 6, 2017 <http://www.pewresearch.org/fact-tank/2017/06/28/10-facts-about-smartphones/>. 2017
21. Muench F, Weiss RA, Kuerbis A, Morgenstern J. Developing a theory driven text messaging intervention for addiction care with user driven content. *Psychol Addict Behav*. 2013; 27(1):315. doi: 10.1037/a0029963 [PubMed: 22963375]
22. Obermayer JL, Riley WT, Asif O, Jean-Mary J. College smoking-cessation using cell phone text messaging. *J Am Coll Health*. 2004; 53(2):71–8. DOI: 10.3200/JACH.53.2.71-78 [PubMed: 15495883]
23. Ybarra ML, Holtrop JS, Prescott TL, Rahbar MH, Strong D. Pilot RCT results of stop my smoking USA: a text messaging-based smoking cessation program for young adults. *Nicotine Tob Res*. 2013; 15(8):1388–1399. DOI: 10.1093/ntr/nts339 [PubMed: 23348969]
24. Kong G, Ells DM, Camenga DR, Krishnan-Sarin S. Text messaging-based smoking cessation intervention: A narrative review. *Addict Behav*. 2014; 39(5):907–917. DOI: 10.1016/j.addbeh.2013.11.024 [PubMed: 24462528]
25. Latimer AE, Krishnan-Sarin S, Cavallo DA, Duhig A, Salovey P, O'Malley SA. Targeted smoking cessation messages for adolescents. *J Adolescent Health*. 2012; 50(1):47–53. DOI: 10.1016/j.jadohealth.2011.04.013
26. Toll BA, Salovey P, O'Malley SS, Mazure CM, Latimer A, McKee SA. Message framing for smoking cessation: The interaction of risk perceptions and gender. *Nicotine Tob Res*. 2008; 10(1):195–200. DOI: 10.1080/14622200701767803 [PubMed: 18188760]

27. Biener L, McCallum-Keeler G, Nyman AL. Adults' response to Massachusetts anti-tobacco television advertisements: impact of viewer and advertisement characteristics. *Tobacco Control*. 2000; 9(4):401–407. DOI: 10.1136/tc.9.4.401 [PubMed: 11106710]
28. Steward WT, Schneider TR, Pizarro J, Salovey P. Need for cognition moderates responses to framed smoking-cessation messages. *J Appl Soc Psychol*. 2003; 33(12):2439–2464. DOI: 10.1111/j.1559-1816.2003.tb02775.x
29. Wilson DK, Wallston KA, King JE. Effects of contract framing, motivation to quit, and self-efficacy on smoking reduction. *J Appl Soc Psychol*. 1990; 20(7):531–547. DOI: 10.1111/j.1559-1816.1990.tb00426.x
30. Wilson DK, Purdon SE, Wallston KA. Compliance to health recommendations: A theoretical overview of message framing. *Health Educ Res*. 1988; 3(2):161–171. DOI: 10.1093/her/3.2.161
31. O'Keefe DJ, Jensen JD. The relative persuasiveness of gain-framed loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review. *J Health Commun*. 2007; 12(7):623–644. DOI: 10.1080/10810730701615198 [PubMed: 17934940]
32. Lee S, Lang A. Redefining media content and structure in terms of available resources toward a dynamic human-centric theory of communication. *Commun Res*. 2015; 42(5):599–625. DOI: 10.1177/0093650213488416
33. Ling, R. *The sociolinguistics of SMS: An analysis of SMS use by a random sample of Norwegians In Mobile communications*. London: Springer; 2005.
34. Brochet F, Naranjo P, Yu G. Causes and consequences of linguistic complexity in non-US firm conference calls. *Harvard Business School Working Paper*. 2012; doi: 10.2139/ssrn.2154948
35. Barnett T, Hoang H, Furlan A. An analysis of the readability characteristics of oral health information literature available to the public in Tasmania, Australia. *BMC Oral Health*. 2016; 16(1):1.doi: 10.1186/s12903-016-0196-x [PubMed: 26745890]
36. Vidrine JI, Simmons VN, Brandon TH. Construction of Smoking-Relevant Risk Perceptions Among College Students: The Influence of Need for Cognition and Message Content. *J Applied Soc Psychol*. 2007; 37(1):91–114. DOI: 10.1111/j.0021-9029.2007.00149.x
37. Reilly, J., Seibert, L. *Language and emotion Handbook of affective sciences*, ed by RJ Davidson, KR Scherer and HH Goldsmith. Oxford: Oxford University Press; 2003. p. 535-59.
38. Shimanoff SB. Expressing emotions in words: Verbal patterns of interaction. *J Commun*. 1985; 35(3):16–31. DOI: 10.1111/j.1460-2466.1985.tb02445.x
39. Shimanoff SB. Types of emotional disclosures and request compliance between spouses. *Commun Monographs*. 1987; 54(1):85–100. DOI: 10.1080/03637758709390217
40. Carey J. Paralanguage in computer mediated communication. *Proceedings of the 18th annual meeting on Association for Computational Linguistics*. 1980; doi: 10.3115/981436.981458
41. Biener L. Anti-tobacco advertisements by Massachusetts and Philip Morris: what teenagers think. *Tob Control*. 2002; 11(suppl 2):ii43–ii46. DOI: 10.1136/tc.11.suppl\_2.ii43 [PubMed: 12034981]
42. Biener L, Ji M, Gilpin EA, Albers AB. The impact of emotional tone, message, and broadcast parameters in youth anti-smoking advertisements. *J Health Commun*. 2004; 9(3):259–274. DOI: 10.1080/10810730490447084 [PubMed: 15360037]
43. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. *J Lang Soc Psychol*. 2010; 29(1):24–54. DOI: 10.1177/0261927X09351676
44. Lieberman MA. Effects of disease and leader type on moderators in online support groups. *Comput Hum Behav*. 2008; 24(5):2446–2455. DOI: 10.1016/j.chb.2008.02.018
45. Pennebaker, JW., Chung, CK., Ireland, M., et al. The development and psychometric properties of LIWC2007. <https://repositories.lib.utexas.edu/handle/2152/31333.2007>
46. Prokhorov AV, Machado TC, Calabro KS, Vanderwater EA, Vidrine DJ, Pasch KP, et al. Developing mobile phone text messages for tobacco risk communication among college students: a mixed methods study. *BMC Public Health*. 2017; 7(1):137.doi: 10.1186/s12889-017-4027-z
47. Bantum EOC, Owen JE. Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychol Assessment*. 2009; 21(1):79.doi: 10.1037/a0014643
48. LIWC. Comparing LIWC2015 and LIWC2007. <http://liwc.wpengine.com/compare-dictionaries/>. 24 March 2016

49. Siu AL. Behavioral and pharmacotherapy interventions for tobacco smoking cessation in adults, including pregnant women: US Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2015; 163(8):622–634. DOI: 10.7326/M15-2023 [PubMed: 26389730]
50. Walker N, Johnston V, Glover M, et al. Effect of a family-centered, secondhand smoke intervention to reduce respiratory illness in Indigenous infants in Australia and New Zealand: A randomized controlled trial. *Nicotine Tob Res.* 2015; 17(1):48–57. DOI: 10.1093/ntr/ntu128 [PubMed: 25156527]
51. Halpern SD, French B, Small DS, et al. Randomized trial of four financial-incentive programs for smoking cessation. *New Engl J Med.* 2015; 372(22):2108–2117. DOI: 10.1056/NEJMoa1414293 [PubMed: 25970009]
52. Arguello J, Butler BS, Joyce E, Kraut R, Ling KS, Rosé C, Wang X. Talk to me: foundations for successful individual-group interactions in online communities. Proceedings of the SIGCHI conference on Human Factors in computing systems. 2006
53. Moseley, DV. Assessment of spelling and related aspects of written expression. In: Beech, JR., Singleton, C., editors. *The psychological assessment of reading.* New York, NY: Routededge; 1997. p. 204-223.
54. Rosselli F, Skelly JJ, Mackie DM. Processing rational and emotional messages: The cognitive and affective mediation of persuasion. *J Exp Soc Psychol.* 1995; 31(2):163–190. DOI: 10.1006/jesp.1995.1008
55. Cliff, N. *Analyzing multivariate data.* Harcourt Brace Jovanovich; 1987.
56. Myers, JL. *Fundamentals of experimental design.* Boston: Allyn and Bacon; 1972.
57. Winer, BJ. *Statistical principles in experimental design.* New York: McGraw-Hill; 1962.



**Figure 1.** Scatter Plots of (a) word count versus word length and (b) affect versus cognitive processing.

*Note.* (a) Stratifies between messages designed to be simple and those designed to be complex, and (b) stratifies between messages designed to be rational and those designed to be emotional.

<sup>a</sup>Value indicates median of variable.



**Table 1**

Examples of text messages as designed by writers

	Gain-framed		Loss-framed	
	Simple	Complex	Simple	Complex
<b>Emotional</b>	Yummy, pie! Nonsmokers can appreciate every single bite of homemade apple pie since the nicotine in cigs hasn't ~ messed up their taste buds! :P	Avoiding cigarettes prevents halitosis. Nonsmokers are not exposed to the disgusting sulfur compounds ~ in tobacco that cause putrid, chronic morning breath! :)	Mike had a hot date on Friday but wouldn't stop smoking cigs with his pals. ~ Now he has rotting yellow teeth & an imaginary date. :(	Devastating news! Smoking 'light' cigarettes will not protect the body from toxicity. ~ All cigarettes rip away approximately a decade from a smoker's lifespan :(
<b>Rational</b>	With 7000+ toxic chemicals in cig smoke, the chances of a nonsmoker getting cancer is really low. Why? ~ They aren't exposed to 60+ cancer-causing chemicals.	Were you aware? Avoiding exposure to chemicals in tobacco smoke can prevent premature skin aging. ~ Nonsmokers maintain skin elasticity by avoiding cigarettes.	Because smoking cigs raises ppl's risk of a heart attack, using it is a bad health choice.	Annually, 16 million Americans have at least one severe disease due to smoking. ~ Smoking any cigarette, even 'lights', leads to difficulties maintaining body condition.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Descriptive statistics for variables of interest after LIWC procedure (N=976)

Variables	Gain-framed	Loss-framed	Total	F	p	n <sup>2</sup>
<b>Message depth</b>						
<b>Complex</b>						
Word count <sup>a</sup>	23.30 (4.15)	22.60 (4.33)	22.88 (4.43)	6.51	0.011	0.010
Word length	28.33 (12.04)	28.31 (12.13)	28.23 (12.16)	<0.001	0.982	<0.001
<b>Message appeal</b>						
<b>Emotional</b>						
Affect	12.45 (7.48)	8.83 (6.26)	10.61 (7.14)	66.78	<0.001	0.064
Positive emotions	5.66 (5.47)	1.48 (2.69)	3.56 (4.79)	229.56	<0.001	0.191
Negative emotions	6.73 (5.53)	7.28 (5.71)	6.98 (5.63)	2.38	0.123	0.002
Anxiety/Fear	3.89 (3.66)	1.87 (2.82)	2.87 (3.42)	92.77	<0.001	0.087
Anger	0.73 (1.81)	0.86 (2.03)	0.79 (1.92)	1.11	0.292	0.001
<b>Rational</b>						
Cognitive processing	10.12 (7.01)	10.70 (7.53)	10.39 (7.29)	1.49	0.22	0.001
Quantity	1.92 (2.91)	1.84 (2.77)	1.88 (2.84)	0.20	0.65	<0.001

*Note.*<sup>a</sup>Word count is the only variable that is not a frequency or ratio.

**Table 3**

Logistic regression model predicting message categorization by writers based on depth (N = 976)

	Messages designed to be simple		
	OR (SE) <sup>a</sup>	95% CI <sup>b</sup>	p
Word count	0.62 (0.02)	0.57-0.66	<0.001
Word length	0.77 (0.01)	0.75-0.80	<0.001
Gain-framed	1.21 (0.26)	0.79-1.86	0.376
Designed to be emotional versus rational	1.78 (0.47)	1.06-3.00	0.029
Affect	1.04 (0.02)	1.00-1.08	0.026
Cognitive processing	1.02 (0.01)	0.99-1.05	0.171
New and emerging product	0.92 (0.19)	0.61-1.39	0.695
Model $\chi^2=744.52$			<0.001

*Note.*

<sup>a</sup>Indicates odds ratio followed by standard error in parentheses.

<sup>b</sup>Indicates 95% confidence interval.

**Table 4**

Analysis of variance comparing message categories based on appeal (N=976)

	Rational	Emotional	F	p	$\eta^2$
<b>Emotion-related variables</b>					
Affect	7.35 (6.27) <sup>a</sup>	13.94 (6.38)	265.05	<0.001	0.21
Positive emotion	1.67 (2.96)	5.47 (5.47)	182.40	<0.001	0.16
Negative emotion	5.61 (5.45)	8.40 (5.45)	64.01	<0.001	0.06
Anxiety and fear	2.49 (3.05)	3.27 (3.72)	12.66	<0.001	0.01
Anger	0.43 (1.50)	1.15 (2.21)	36.13	<0.001	0.03
Sad	0.88 (2.13)	1.05 (2.34)	1.42	0.233	<0.01
<b>Cognitive-related variables</b>					
Cognitive processing	11.70 (7.94)	9.14 (6.30)	31.24	<0.001	0.03
Quantity	2.16 (3.14)	1.61 (2.48)	9.31	0.002	<0.01

<sup>a</sup>Values show mean and standard deviation.

**Table 5**

Logistic regression analysis predicting messages designed by writers based on appeal (N=976)

	Messages designed to be emotional		
	OR (SE) <sup>a</sup>	95% CI <sup>b</sup>	p
Affect	1.23 (0.02)	1.19-1.27	<0.001
Cognitive processing	0.96 (0.01)	0.94-0.99	0.004
Gain- versus loss-framed	0.38 (0.07)	0.27-0.54	<0.001
Designed to be simple versus complex	1.77 (0.44)	1.08-2.89	0.023
Word count	1.37 (0.01)	1.29-1.45	<0.001
Word length	1.01 (0.01)	0.99-1.03	0.137
New and emerging product	1.47 (0.25)	1.05-2.05	0.025
Model $\chi^2=461.49$			<0.001

<sup>a</sup>Indicates odds ratio followed by standard error in parentheses.<sup>b</sup>Indicates 95% confidence interval.

Table 6

Identified themes for nicotine and tobacco-risk communication in messages from the Texas-TCORS Library (N=976)

Themes	N(%) <sup>a</sup>	M(SD) <sup>b</sup>	Examples of Messages
Health	734 (75)	6.06 (5.06)	'Ppl who don't use hookah help keep their <i>lungs</i> safe from <i>infection</i> since hookah pipes can be filled with fungi.' 'Don't forget: smoking cigs while pregnant isn't the only way to raise a baby's risk of <i>asthma</i> . Another way to raise the baby's risk is to be around secondhand smoke.'
Body	373 (38.22)	2.54 (3.72)	'Hooray for healthy <i>brains</i> ! Stronger <i>disease-fighting systems</i> protect ppl who don't smoke cigs from <i>brain</i> -swelling infections. Go <i>brain</i> power! :)' 'Users have to worry about e-juice spills on the <i>skin</i> which can cause nasty <i>heart</i> problems. Using e-cigs means more risk of being poisoned! :('
Work and Marketing	275 (28.18)	1.74 (3.32)	' <i>Consumers</i> protect their health by ignoring attempts from tobacco <i>companies</i> to target them with concepts of "lighter" &/or "more natural" <i>products</i> .' 'Tobacco users lower their chances of getting a <i>job</i> . In Texas, there aren't any <i>laws</i> to protect them from being denied <i>work</i> just because they use tobacco.'
Leisure	255 (26.05)	1.82 (3.60)	'Social trends- here today, gone tomorrow but their effects can last much longer. Ppl help themselves by avoiding hookah <i>bars &amp; hanging out</i> in smoke-free places.' 'Wondering why some college kids avoid hookah <i>bars</i> ? They know that hookah is bad & are trying not to infect their lungs.'
Money	118 (12.05)	0.65 (1.88)	'The main cause of lung disease is cig smoke. Each year lung disease causes millions of deaths, high health <i>costs &amp; work-force-losses</i> .' 'Alternative" cigarettes, like bidis, are <i>marketed</i> as additive-free. Nonsmokers <i>benefit</i> by avoiding all cigarettes including "alternative" <i>products</i> .'
Death	34 (3.45)	0.16 (0.86)	'Cigar use is estimated to cause approximately 9,000 premature <i>deaths</i> every year. That is almost 140,000 years of potential <i>life lost</i> annually.' 'Don't be a loser at the game of life! The nicotine in e-juice refills can be <i>deadly</i> , so drinking it can lead to a <i>fatal</i> overdose. Play it safe! :('
Sexuality	33 (3.38)	0.17 (0.96)	'Snus, tobacco in a pouch held inside the mouth, poses greater risk to <i>pregnancies</i> . Pregnant women who use snus increase their risk of a preterm <i>birth</i> .' 'Unlike nonsmokers, male smokers increase their risk of <i>erectile</i> dysfunction by 30%. Smoking cigarettes increases the risk of developing this disorder.'
Religion	22 (2.25)	0.09 (0.63)	'Using electronic cigarettes raises the intention of using conventional cigarettes because users create a <i>demon</i> of addiction that some subdue with conventional cigarettes. :('
			'Avoiding the scourge of secondhand smoke is a <i>blessing</i> for pregnant women! Mothers who avoid the toxic smoke of cigarettes are able to enjoy a healthier and happier baby. :)'

<sup>a</sup>Indicates sample size and percentage of text messages identified under the theme.

<sup>b</sup>Indicates mean frequency of words depicting the theme and standard deviation. Words in italics are identified by LIWC under the theme in question.