

Understanding Socio-Cultural Factors related to Obesity: Sentiment Analysis on related Tweets

Albert Park¹, Delia S. West²

¹University of North Carolina at Charlotte, Charlotte, North Carolina, United States, ²University of South Carolina, Columbia, South Carolina, United States

Objective

We aim to better understand socio-cultural factors (i.e., attitudes or perceptions of cultural groups) associated with food consumption and weight loss via sentiment analysis on tweets, short messages from Twitter.

Introduction

Obesity can lead to the death of at least 2.8 million people each year [1], yet the rate of obesity around the world has continuously increased over the past 30 years [1]. Societal changes, including increased food consumption and decreased physical activity, have been determined as two of the main drivers behind the current obesity pandemic [2]. Examining socio-cultural factors (i.e., attitudes or perceptions of cultural groups) [3] associated with food consumption and weight loss can provide important insights to guide effective interventions and a novel surveillance approach to characterize population obesity trends from sociological perspectives. The primary goal of this study is to examine socio-cultural factors associated with food consumption and weight loss by conducting sentiment analysis on related online chatters. The secondary goal is to discuss the potential implications of being exposed to these different chatters in the online environment.

Scientific evidence in support of using social media to understand socio-cultural factors and its potential implications can be illustrated in two concise assertions. First, online chatters, including discussions on social media, have been shown to be an effective data source for understanding public interests [4,5]. Second, prolonged participation in social media has been suggested to have impacts on users [6-8].

Methods

In this study, we examined Twitter (www.twitter.com), a highly popular, free-to-use, micro-blogging social media platform that can instantly broadcast short messages to the world. These short messages are called Tweets and we collected weight loss related and food consumption related Tweets using Python library called *Tweepy* [9]. We used hashtags from a previous study [10], including *#weightloss*, *#diet*, *#fitness*, and *#health* for collecting weight loss related tweets. Similarly, we used *#Food*, *#FoodPorn*, and *#Foodie* to collect food consumption related tweets. We then used a rule-based model called Vader [11], a sentiment analysis tool (i.e., computational process of categorizing sentiment) developed for social media text, to measure tweets' sentiment. We used the compound score, which is a normalized and weighted composite score that ranges from -1.0 (most negative) to 1.0 (most positive). Lastly, we conducted independent sample *t*-test to compare the sentiments of two types of tweets.

Results

We collected 81,535 (from 41,436 unique user ID) weight loss related tweets from August 30th to September 2nd of 2018 and 86,277 (from 36,977 unique user ID) food consumption related Tweets from August 28th to September 2nd of 2018. The mean sentiment score for weight loss related tweets was 0.17 (sample standard deviation: 0.39), whereas the mean sentiment for food consumption related Tweets was more positive, scoring 0.26 with sample standard deviation of 0.34. The independent sample *t*-test suggests that the sentiment difference between the two types of tweets is statistically significant ($t=52.10$, $p < .001$). However, it is important note that the mean sentiment for both types of tweets was in the positive range.



ISDS Annual Conference Proceedings 2019. This is an Open Access article distributed under the terms of the Creative Commons AttributionNoncommercial 4.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0/>), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conclusions

We present preliminary findings concerning socio-cultural factors associated with food consumption and weight loss within twitter chatters. Our initial results suggest that individuals expressed more positive sentiment when tweeting about food consumption than when tweeting about weight loss. The results not only reflect the social norms of social media, Twitter in this particular study, but also suggest how social media can indirectly promote more food consumption over weight loss via social norms theory [12] and how online social norms can reach individual members. This is especially important for young adults, the main demographic user group for social media [13], as they develop lasting health related habits and behaviors. Although in its infancy, our research suggests that online sociocultural environment could be a potential socio-environmental risk factor for obesity. The next step is to utilize the findings to create online sociocultural environment that can promote the healthy choices.

Acknowledgement

We restricted our analysis to publicly available discussion content. The study was determined to be not human subjects by the University of North Carolina-Charlotte's Institutional Review Board (Ethics Committee). Author AP's contribution to this research was partially supported by AP's startup research funding provided by the Department of Software and Information Systems, University of North Carolina-Charlotte. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agency.

References

1. World Health Organization. Global Health Observatory (GHO) data: Obesity [Internet]. 2009. Available from: http://www.who.int/gho/ncd/risk_factors/obesity_text/en/
Archived at: <http://www.webcitation.org/6rQICh7Oq>
2. Hill JO. 1998. Environmental Contributions to the Obesity Epidemic. *Science*. 280(5368), 1371-74. [PubMed](#)
<https://doi.org/10.1126/science.280.5368.1371>
3. Swinburn B, Egger G, Raza F. 1999. Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity. *Prev Med*. 29(6), 563-70. [PubMed](#) <https://doi.org/10.1006/pmed.1999.0585>
4. Park A, Conway M. 2017. Towards Tracking Opium Related Discussions in Social Media. *Online J Public Health Inform*. 9(1), e073. <https://doi.org/10.5210/ojphi.v9i1.7652>
5. Park A, Conway M. Tracking Health Related Discussions on Reddit for Public Health Applications. Annu Symp proceedings AMIA Symp. 2017;2017:1362-71.
6. Park A, Conway M. 2018. Harnessing Reddit to Understand the Written-Communication Challenges Experienced by Individuals With Mental Health Disorders: Analysis of Texts From Mental Health Communities. *J Med Internet Res*. 20(4), e121. [PubMed](#) <https://doi.org/10.2196/jmir.8219>
7. Park A, Conway M. 2017. Longitudinal Changes in Psychological States in Online Health Community Members: Understanding the Long-Term Effects of Participating in an Online Depression Community. *J Med Internet Res*. 19(3), e71. [PubMed](#) <https://doi.org/10.2196/jmir.6826>
8. Park A, Hartzler AL, Huh J, McDonald DW, Pratt W. Homophily of Vocabulary Usage: Beneficial Effects of Vocabulary Similarity on Online Health Communities Participation. Annu Symp proceedings AMIA Symp. 2015;2015:1024-33.
9. Roesslein J. Tweepy [Internet]. 2009. Available from: <http://www.tweepy.org>
10. Turner-McGrievy GM, Beets MW. 2015. Tweet for health: using an online social network to examine temporal trends in weight loss-related posts. *Transl Behav Med*. 5(2), 160-66. [PubMed](#)
<https://doi.org/10.1007/s13142-015-0308-1>
11. Hutto CJ, Gilbert E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. Eighth Int Conf Weblogs Soc Media (ICWSM-14). 2014;216-25.



ISDS Annual Conference Proceedings 2019. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0/>), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

12. Perkins HW, Berkowitz AD. 1986. Perceiving the community norms of alcohol use among students: some research implications for campus alcohol education programming. *Int J Addict.* 21(9–10), 961-76. [PubMed
https://doi.org/10.3109/10826088609077249](https://doi.org/10.3109/10826088609077249)
13. Smith A, Anderson M. Social Media Use in 2018 [Internet]. Pew Research Center Internet, Science & Tech. 2018. Available from: <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
Archived at: <http://www.webcitation.org/70NIId9Vv2>

Table 1. Characteristics of the analyzed datasets

	Weight Loss Tweets	Food Consumption Tweets
Dates	August 30, 2018 - September 2, 2018	August 28, 2018 - September 2, 2018
Number of Tweets	81,535	86,277
Number of Unique User ID	41,436	36,977
Mean Sentiment Score (sample standard deviation)	0.17(0.39)	0.26(0.34)

The independent sample t-test suggests that the sentiment difference between the two types of tweets is statistically significant ($t=52.10$, $p < .001$)

