HHS Public Access

Author manuscript

Drug Alcohol Depend. Author manuscript; available in PMC 2019 December 01.

Published in final edited form as:

Drug Alcohol Depend. 2018 December 01; 193: 75–82. doi:10.1016/j.drugalcdep.2018.08.021.

#drunktwitter: Examining the relations between alcohol-related Twitter content and alcohol willingness and use among underage young adults

Dana M. Litt[a], **Melissa A. Lewis**[a], **Emma S. Spiro**[b], **Lovenoor Aulck**[b], **Katja A. Waldron**[c], **Maya K. Head-Corliss**[d], and **Alex Swanson**[e]

[a]Department of Health Behavior and Health Systems, University of North Texas Health Science Center, Fort Worth, TX

[b]Information School, University of Washington, Seattle, WA

[c]Department of Biobehavioral Health, Pennsylvania State University, State College, PA

[d]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA

[e]Department of Educational, School, and Counseling Psychology, University of Missouri, Columbia, MO

Abstract

Purpose: Despite the importance of social networking sites on young adult alcohol use, few studies have examined Twitter as a conduit for sharing drinking behavior. However, this work generally uses random samples of tweets and thus cannot determine the extent to which Tweets correspond with self-reported drinking cognitions or behaviors. The primary aims of the present study were to (1) document basic patterns of alcohol-related Twitter activity in a subsample of young adult drinkers, and (2) examine whether willingness to drink, alcohol use, and negative consequences are associated with alcohol-related tweeting behavior.

Methods: 186 young adults age 18-20 completed an online survey and provided Twitter handle information. From these participants, a random sample of 5,000 Tweets was coded by a trained team to determine whether tweets were related to alcohol use or not. Ordinary least squares

Correspondence: Dana M. Litt, Department of Health Behaviors and Health Systems, School of Public Health, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, TX 76107, Phone: 817-735-5453, dana.litt@unthsc.edu.

Contributors

Dana Litt was involved with study conceptualization, paper writing, and coordination of authors.

Melissa Lewis was the Principal Investigator from which the study data was drawn and assisted the first author with manuscript preparation. Emma Spiro and Lovenoor Aulck were responsible for the collection and analysis of Twitter data. Katja Waldron, Maya Head-Corliss, and Alex Swanson were responsible for assisting with manuscript preparation, and in particular writing the methods section. All authors conducted multiple rounds of edits and revisions prior to submission. All authors read and approved the final manuscript.

Conflict of Interest

No conflict declared.

regression analyses were conducted to determine whether the proportion of alcohol-related Tweets is associated with self-reported alcohol use willingness, behaviors, and negative consequences.

Results: Results indicated that not only are alcohol-related tweets common among young adults, but that the proportion of one's overall tweets that are related to alcohol is significantly associated with willingness to drink, alcohol use, and negative consequences.

Conclusions: The results of this study are an important step to understanding how digital behavior (e.g., posting about alcohol on Twitter) is related to an individual's self-reported drinking cognitions, alcohol use, and negative consequences and has implications for the way Twitter data can be used for public health surveillance and interventions.

Keywords

Twitter; Alcohol use; Young adults; Social networking sites; Drinking; Cognitions

1. Introduction

1.1 Social networking sites and young adult alcohol use

It has been argued that social networking sites (SNS) expand young adults' (YAs') exposure to peer risk behavior, including alcohol use. When viewing others' profiles, users can learn about others' interests, hobbies, social lives, and health and risk behaviors (Ellison et al., 2007). Given how much time YAs spend on SNS in conjunction with the multiple ways (e.g., text, photo) and opportunities to communicate about alcohol, SNS are an influential risk-conducive environment for YA alcohol use (McCreanor et al., 2013; Moreno and Whitehill, 2014). With many SNS profiles among YAs in the United States including alcohol-related content or displays, and the majority of displays being pro-alcohol or favorable toward heavy alcohol use (Cavazos-Rehg et al., 2015; Moreno et al., 2012, 2016) research indicates that YAs both produce and are exposed to alcohol-related content via SNS. Both experimental and longitudinal studies have demonstrated that viewing SNS alcohol displays is significantly associated with increased risky drinking cognitions, alcohol use, and negative consequences (Boyle et al., 2016; Hoffman et al., 2017; Huang et al., 2014; Geusens and Beullens, 2017; Litt and Stock, 2011; Nesi et al., 2017; Tucker et al., 2013). In addition, several studies have reported positive associations between the sharing of alcohol-related content and alcohol consumption (D'Angelo et al., 2014; Geusens and Beullens, 2016; Moreno et al., 2015; Nesi et al., 2017). Research has suggested that posting alcohol-related content may be a direct reflection of the sharer's alcohol use (D'Angelo et al., 2014; Geusens and Buellens, 2016; Westgate et al., 2014) and may also be a reflection of alcohol-related social cognitions (i.e., perceived norms and attitudes), which are known to predict alcohol use (D'Angelo et al., 2014; Westgate et al., 2014). Notably, research indicates that older adolescents' heavy episodic drinking behavior during the preceding year predicted the frequency of which they posted alcohol references on SNS (Geusens and Buellens, 2017) and that frequent posting of alcohol use content on SNS predicts later alcohol use (Erevik et al., 2017). Further, several studies have linked alcohol-related cognitions and alcohol postings on social media (D'Angelo et al., 2014; Erevik et al., 2017; Geusens and Buellens, 2016a; Guesens and Buellens, 2016b; Stoddard et al., 2012; Westgate et al., 2014).

It is important that research continue to determine whether there is a relationship between one's own SNS behavior and self-reported alcohol-related risk cognitions. Therefore, a growing literature of experimental and longitudinal studies shows that SNS alcohol-related displays are common among YAs, and these displays are associated with future problematic drinking or vice versa. However, it is important to note two major limitations of this field. First, the majority of research on SNS and alcohol use has focused on Facebook (D'Angelo et al., 2014; Moreno et al., 2015; Moreno et al., 2012) and has used self-reports of alcohol-related SNS displays (Erevik et al., 2017; Geusens and Beullens, 2016; Stoddard et al., 2012). Secondly, this manuscript references a snapshot of tweets that may not overlap temporally with the survey, meaning that the tweets may have been posted both before and after the survey was completed by participants.

1.2 Twitter as a prominent social networking site

There are over 11.7 million active users between the ages of 18-24 on Twitter, comprising approximately 45% of all young adults in that age bracket (Pew, 2018). In addition, roughly 46% of all Twitter users visit the site at least daily (Pew, 2018). Two important differences between Facebook and Twitter are that there is less interaction and oversight from older adults/parents on Twitter (23%) than on Facebook (74%) (Duggan et al., 2015) and the majority (88.2%) of Twitter content and user profiles are publicly visible to all internet users, while on Facebook most information is shared only with selected individuals. It is important to note that personal disclosures are less common on Twitter than on other SNS, with users instead often exchanging information like news, and that passive consumption is more common (Huberman et al., 2008; Kwak et al., 2010; Phua et al., 2017). However, although research indicates that Facebook users were more likely to be engaged through a large social network and daily log-ins than Twitter users, participants who chose to display alcohol references on Twitter were more likely to use the site daily (Moreno et al., 2012). Because Twitter profiles are more likely to be public than other forms of social media, there may be greater potential to incorporate Twitter content into intervention efforts. Interventions that involve private profiles generally entail researchers gaining access to private profiles. With Twitter, there is the unique option of providing intervention content to public profiles.

Twitter users maintain many types of social relationships, including subscription ties called "following" relationships. Following relationships form the basis for the underlying social network on Twitter, facilitating information diffusion and potential peer influence. Following relationships represent subscriptions to other users' content, allowing tweets to be automatically delivered (and therefore readable) from one user to another. Tweets are publicly visible by default but users can restrict message delivery to followers only. In addition, Twitter users have developed platform-specific conventions to organize content. Hashtags (indicated by the # symbol) are used to denote topical channels, and to facilitate search and navigation of content. Users can also "mention" or reply to specific accounts using @username. The Twitter platform facilitates re-sharing of content through re-posting features, called retweets. Retweeting a message allows users to repost a message from another Twitter account and share it with one's own followers. Emojis, or pictographic images (Marengo et al., 2017), are increasingly used as forms of SNS communication, especially among YAs, and appear to serve as a surrogate of non-verbal cues, contributing to the overall meaning of written messages (Marengo et al., 2017).

1.3 Twitter as a risk conducive environment

Despite the argued importance of social networking sites on alcohol use among YAs, few studies have examined Twitter as a social media conduit for sharing drinking behaviors (Cavazos-Rehg et al., 2015; West et al., 2012; Moreno et al., 2016). Research indicates that young adults who choose to display alcohol references on Twitter are more likely to be daily Twitter users and the number of alcohol-related tweets is associated with number of drinks (Moreno et al., 2016). However, this study did not take into account the relative proportion of alcohol-related tweets relative to overall tweets, evaluate cognitions or negative consequences as outcomes and was comprised of only college students who reported using both Facebook and Twitter, thus reducing generalizability. In addition, the majority of research examining Twitter and alcohol content focuses on data collected using convenience samples of public tweets and general temporal trends in alcohol-related tweets (West et al., 2012), underage YA access to alcohol-related tweets, or frequency of alcohol-related tweet content (Cavazos-Rehg et al., 2015). Specifically, West and colleagues (2012) examined the temporal trends of drinking-related tweets generated from Twitter users and found that an increase in pro-drinking tweets (i.e., related words to "drunk" and intoxication) occurred during nights (between 10pm and 2am) and weekends (Friday and Saturday nights). In addition, Cavazos-Rehg and colleagues (2015) demonstrated that there was a significant amount of alcohol-related Twitter "chatter" over the course of one month (about 400,000 drinking-related tweets per day) and that there are more positive views toward drinking (pro-drinking) rather than tweets that portray drinking in a negative manner (anti-drinking). Despite the frequent use of emojis, most research to date has focused on emojis as an indicator of emotions and personality (Marengo et al., 2017) and has not looked at emojis as a specific form of alcohol-related Twitter chatter.

One major limitation of most work so far examining Twitter is that researchers use a random sample of Tweets and thus are not able to determine the extent to which these Tweets correspond with self-reported drinking cognitions or behaviors of individuals, which has been listed as a major limitation of this body of work (Krauss et al., 2017). While these studies are important and provide descriptions of alcohol behavior as expressed on Twitter, they are limited in their generalizability because next to nothing is known about the characteristics (cognitions or behavior) of the individuals whose posts are being analyzed.

Therefore, the primary aims of the present study are to (1) document basic patterns of alcohol-related Twitter activity in a subsample of YA drinkers, and (2) examine whether self-reported willingness/openness to use alcohol (Gerrard et al., 2008), and alcohol use are associated with the proportion of one's overall tweets that are alcohol-related.

2. Materials and methods

2.1 Participants and procedures

Data for this study was collected as part of a larger NIH-funded study (R01AA021379) that investigated health risk behaviors among young adults. Participants for this study were YAs in the United States aged 18 to 20 who were recruited nationally through various methods including paid online recruiting (i.e., Facebook, Twitter, and Craiglist ads in large US cities), advertisements in-print, and in-person recruiting. A total of 1,145 participants responded to the study advertisements and were invited to the study survey. Of those, 1002 (88%) completed the baseline survey. Participants were informed they would receive a $25 gift certificate for completing the 30-45 minute online survey and would also be entered in a drawing to win an Apple iPad or $100 gift card. A federal certificate of confidentiality was issued by the National Institutes of Health (NIH) to protect identifiable research information from forced disclosure. All study procedures were approved by the University of Washington's institutional review board, and no adverse events were reported.

At the conclusion of the study, all participants were emailed regarding our intentions with utilizing the Twitter data and participants were given a chance to opt-out of our Twitter collection procedures if they had previously provided us with their Twitter information. Any participants who wished for their Twitter data to be excluded were removed from the data collection procedures immediately. Of the 1,058 participants who completed the baseline survey, 406 (38.4%) indicated that they were Twitter users. Of all Twitter users, 209 (17.6% of all; 45.8% of Twitter users) provided usable Twitter data . For Twitter data to be usable, participants must have provided a valid (i.e., existing) Twitter handle for a public Twitter account. In addition, participants must have been willing to allow their Twitter data to be subsequently used in the study, following an IRB-approved opt-out follow-up letter after completion of the survey. Twitter accounts were subsequently checked for obvious misreporting by participants (e.g., spam accounts) and whether participants had provided multiple accounts. In total, 8 accounts were eliminated, leaving 201 valid accounts used in the subsequent analysis, of which 186 had Twitter profiles from which data could be gathered. For the approximately 16% of participants (N=186) whose Twitter data was obtained, the Twitter REST Application Programming Interface (API) was used to collect social media posts (i.e., tweets) from participants based on usernames provided in the survey. The Twitter API provides access to a maximum of 3,200 historical posts per user, spanning a variable period of time depending on how active the specific user is within the platform. Twitter data collection began in December 2014 and all tweets for every user until June 2015 were collected. In addition, using the API, tweets for some participants were collected as far back as May 2008. The timestamps for the tweets spanned anywhere between 1 to 2343 days (approx. 6.4 years) for each user with a median time between first and last tweets of 908 days (approx. 2.5 years) across users. The variability between data collection periods across users was a result of frequency of tweeting, with respect to both historical activity and how frequently each user tweeted between December 2014 and June 2015. Because Twitter counts were normalized on a per-tweet basis as described below, the varying timespans for tweets were used to allow for a maximum amount of data to be analyzed on a per-user basis.

A list of drinking-related terms was developed based on input from our research team, web searches, and searching urbandictionary.com, a free online resource that tracks modem slang. Once the list of initial drinking-related words were developed, investigators scanned a random sample of 1,000 tweets in order to determine frequency of alcohol-related words and adapt as needed, resulting in a total of 17 alcohol-related terms, including “drunk,” “wasted,” and “hangover” (see Table 2 for full list of terms). In addition to drinking-related words, a list of drinking-related emojis was also developed, which included emojis such as the “beer mug” and the “wine glass” (see Table 2 for full list of emojis). The Twitter Streaming API was used to track the alcohol-related terms and phrases and custom programmatic scripts were used to access Twitter data; data were subsequently processed, cleaned, and archived.

2.2 Coding procedures

Once the initial dataset was extracted by Twitter API, the content of a random sample of 5,000 Tweets was coded to summarize their main themes, and in particular whether the alcohol content in the tweet was about alcohol or alcohol use, or not alcohol-related. Two members of the research team with expertise in substance use research created initial definitions of what constituted alcohol-related tweets versus not alcohol-related and then scanned 300 random alcohol-related Tweets. Subsequently, a team of three trained coders used a five-step process (Lee et al., 2007): (1) investigators independently reviewed alcohol related tweets; (2) investigators discussed initial categories (alcohol-related vs. not) of responses to determine tentative definitions; (3) definitions were provided via a coding manual to three trained independent undergraduate and staff coders in order to code each tweet based on the content; (4) a random 10% of responses were coded for reliability (i.e., high consistency between raters) and discussion/revision of coding scheme in which we apply a majority rules decision such that agreement by two or more coders determines the categorization if there are disagreements; and (5) when final coding scheme had been determined, remaining responses were coded.

In the analyses that follow we focus on tweets coded as not alcohol related versus alcohol-related. A tweet coded as not alcohol-related required that the coder must have 100% confidence that it is NOT referring to alcohol use (e.g., “I’ll just be here drinking my milk and eating cookies”). A tweet coded as alcohol use is one in which the tweeter is explicitly referring to alcohol use (e.g., “I am experiencing the worst hangover of my life” or “That drunk chick is going to have issues tomorrow”).

Overall level of coder agreement for the sub sample of test Tweets was good. An inter-rater reliability analysis using the Fleiss’ Kappa statistic was performed to ascertain consistency among raters and indicating the following kappas: not Alcohol-Related, $\kappa > 0.83$ and $\kappa > 0.75$ for tweets about alcohol use (all ps < .001)

2.3 Baseline survey measures

2.3.1 Willingness.—To assess drinking willingness, participants were presented with a scenario that involved drinking at a party and rated their willingness to engage in each of five actions (Gerrard et al., 2008). Sample items include “choose a non-alcoholic drink” and "stay and have one more drink" (α = .85). Response options ranged from 0 (not at all willing) to 4 (completely willing).

2.3.2 Drinks per week.

The Daily Drinking Questionnaire (DDQ) (Collins et al., 1985) was used to measure the number of standard drinks consumed on each day of a typical week during the last three months. Weekly drinking was computed by summing the standard number of drinks on a scale from 0 (0 drinks) to 25 (25+ drinks) for each day of the week and summing all days to create a score for the total number of drinks per week.

2.3.3 Problem drinking.

The Alcohol Use Disorders Identification Test (AUDIT) (Babor et al., 2001; Bohn et al., 1995) is a 10-question scale, with most answers on a 0 to 4 Likert scale assessing consumption, dependence, and harm or consequences of alcohol use. Questions include an assessment of the frequency of drinking alcohol (never, monthly or less, 2-4 times a month, 2-3 times a week, and 4 or more times a week) and the frequency of binge drinking (never, less than monthly, monthly, weekly, and daily) as well as the negative consequences associated with alcohol use. The AUDIT scores can theoretically range from 0 to 40; a score of 8 or higher indicates that the person is at risk for problem drinking, (α = .85)

2.3.4 Alcohol-related consequences.

The Young Adult Alcohol Consequences Questionnaire (YAACQ) (Read et al., 2006) was used to determine any negative consequences experienced while drinking. Participants were presented with 48 items and were asked which, if any, consequences they had experienced in the past 30 days. Response options were 1 = *yes* and 0 = *no*. A sum score was calculated to determine the total number of alcohol-related negative consequences.

2.4 Analysis plan

The Twitter REST API was used to collect more than 341,380 posts from the final set of 186 participants spanning the period January 1, 2010 through March 31, 2016. Twitter's API allows researchers to gather information using structured queries as a means of accessing information regarding specific phrases or key words (McCormick et al., 2015). For each Twitter user, the number of alcohol-related twitter displays (including both alcohol-related words and alcohol-related emojis) was normalized by the number of total tweets collected for the user (i.e., on a per-tweet basis) thus yielding a proportion of alcohol-related Twitter displays.

Associations between the proportion of alcohol-related Twitter displays and drinking/ willingness outcomes (as measured at baseline survey) were determined using regressions. In the regression models, each outcome was treated as a dependent variable in isolation with alcohol-related twitter displays included as explanatory variables along with demographic covariates. Preliminary analyses revealed non-normal distributions for drinking outcomes (i.e., drinks per week, consequences, problem drinking), with the variance being substantially greater than the mean. Because these drinking outcomes closely followed a negative binomial probability distribution, a generalized linear modeling approach with the distribution specified as negative binomial (i.e., negative binomial regression) was selected as the primary analysis strategy for drinking outcomes (Atkins and Gallop, 2007). Regression models for willingness to drink treated the response as a continuous variable with a normal distribution. Ordinary least squares linear regression was used to model the association between outcomes and alcohol-related twitter displays. Gender, age, race, and ethnicity were included in all analyses as covariates based on previous associations with alcohol consumption (O'Malley and Johnston, 2002; Read et al., 2002). Dummy variables were created for each demographic variable except age and the data was normalized using min-max normalization prior to running regressions. All data analysis for this project was performed using the R statistical computing platform and/or the Python programming language (R Core Team, 2013).

3. Results

3.1 Differences between Twitter data providers and Twitter data non-providers

Table 1 shows the demographic composition of all participants and across two subgroupings: those who indicated they were active on Twitter with those who were not active on Twitter, as well as those for whom Twitter data was used in this study with those for whom Twitter data was not used in this study, assuming Twitter activity. Of note is the relative consistency in demographic composition across the groups, as indicated by the p-values (calculated using t-tests for age and chi-squared tests for categorical data). For the data set used in the current analyses, those with Twitter profiles from which data could be gathered, the mean age was 19.17 years old ($SD = 0.83$). Gender and racial representation of the sample for whom Twitter data was used was 54.3% female, 12.9% Asian, 21.0% African American, 48.9% White, 0.5% American Indian/Alaska Native, 0.0% Native Hawaiian/Pacific Islander, and 16.7% Other/Mixed. For ethnicity, 14.5% identified as Hispanic, and for schooling, 13.4% were not in any form of school, 60.2% were attending a 4-year university, 18.8% were attending a community college, 2.2% were attending a technical/vocational college, and 2.2% were in high school.

To check for possible selection bias, we compared self-reported willingness to drink, alcohol use, and alcohol-related negative consequences between participants who shared useable Twitter data with those that did not. Results confirmed that there were no significant differences between the study groups on any of the self-reported outcomes, in terms of willingness to drink, behavior, and consequences (all $ps > 0.05$).

3.2 General descriptive information

Of the 186 usable Twitter profiles, we examined tweeting behaviors, along with Twitter specific conventions such as use of hashtags and URLs in tweets. We collected a total of 337,935 tweets, indicating that Twitter users in the dataset posted an average of 1,816 tweets, with a max of 23,085, demonstrating extremely high rates of use. Tweets were on average 12 words long (Twitter imposes a 140 character limit), and contained an average of 0.31 hashtags and 0.43 URLs.

3.3 Alcohol-related content descriptive information

Results from the first round of coding indicated that 58.9% of the tweets extracted (based on a priori terms) contained alcohol content. Examples of tweets that were coded as not being drinking-related are “I can’t stop drinking coffee” or “I dare you to drink a gallon of milk.” These non-alcohol related terms were removed from the dataset for the subsequent analyses.

Table 2 shows the proportion of tweets in the dataset that contain the a priori drinking-related terms, the number of unique users who used each term, the proportion of all users who used each term, the number of tweets that contained each term. The most frequent terms reported were *drunk, drink, drinking,* and *beer.* Further, over half (53%) of study participants posted a tweet containing the word *drunk*, and over one-third (34%) of participants tweeted a message containing the term *wasted.* In addition, over one-fourth (28%) of users posted at least one tweet with an alcohol-related emoji. (See Table 2 for the full list of alcool-related emojis.)

3.4 Relationships between Twitter posts and self-reported cognitions and behavior

3.4.1 Willingness.—None of the demographic variables (gender, race, ethnicity, and age) predicted willingness to drink. However, having a higher proportion of alcohol tweets was positively predictive of greater willingness to drink. (See Table 3.)

3.4.2 Drinks per week.—None of the demographic variables (gender, race, ethnicity, and age) predicted typical drinks per week, but having a higher proportion of alcohol tweets was positively predictive of greater numbers of drinks per week consumed. (See Table 4.)

3.4.3 Alcohol-related negative consequences.—As with drinks per week, the only significant predictor of alcohol-related negative consequences was the proportion of alcohol tweets a user posted. (See Table 4.)

3.4.4 Problem drinking.—Consistent with the results of the other outcomes, the proportion of alcohol-related tweets was the only significant predictor of scores on the AUDIT. (See Table 4)

4. Discussion

Despite limited knowledge, researchers and practitioners aim to use SNS data to infer health behaviors of users, allowing for early detection, intervention, and prevention. However, research has yet to relate SNS expressions of drinking on Twitter to self-reported drinking data. Thus, research to date cannot determine how alcohol-related tweets are associated with alcohol use and how much can be derived about drinking behavior solely from Twitter data. Filling this gap, the proposed research evaluated whether alcohol-related tweets are associated with drinking cognitions, alcohol use and related risks. The results of the present study are an important step to shed some light on how digital behavior (i.e., posting about alcohol on Twitter) is related to an individual’s self-reported drinking cognitions, alcohol use, and experienced negative consequences.

Results of the current study indicate that individuals willing to share their Twitter handles with researchers are not significantly different from those who do not share, indicating that obtaining Twitter handles this way is not likely to introduce bias to the sample. Further, not only are alcohol-related tweets (both text-based and emoji-based) common, but the frequency (as part of a larger proportion of alcohol tweets relative to total tweets) of posting alcohol-related tweets on Twitter is significantly associated with willingness to drink, self-reported alcohol use, and related negative consequences. Our results are in line with the Media Practice Model, which suggests that individuals use media outlets to disclose actual behaviors or behavioral intent and make a statement about identity (Steele, 2005).

5. Health and clinical implications

Data from SNS, such as Twitter, can be used not only to validate current methods that infer health behaviors based solely on social media data, but also to develop novel techniques for estimating prevalence and consequences of alcohol use in specific populations. For example, use of SNS alcohol displays can be used to help identify individuals who meet criteria for being “at risk” based on their Twitter displays and social networks. This could be used in intervention research as the Twitter engagement center allows targeted messages to be sent to individuals based on several criteria, including targeting people who search, tweet about, or engage with specific alcohol-related keywords. Further, the current findings indicated that there were different tenses used in Tweets about alcohol. For example, Tweets with present tense (i.e., “I am so drunk right now”) could allow for in-the-moment text messages to be given to help reduce drinking or risk when needed most. Moreover, texts with past tense (i.e., “I got way too drunk last night!”) could be framed to consider the previous days alcohol use and how it affects goals, etc. Further, because the present study focused on the proportion of overall tweets that contained alcohol references as opposed to a raw frequency of alcohol Tweets, it may help shed light on what frequency of alcohol-related tweeting is problematic relative to overall patterns of tweeting behavior. Understanding how one’s alcohol-related tweeting behavior fits into their overall Twitter behavior may provide valuable insight into the role that alcohol use plays in an individual’s social media identity and provide avenues for understanding which individuals are the most in need of intervention based on alcohol-related Twitter behavior. Because Twitter has a higher percentage of public profiles than other social networking sites, therefore it might be more feasible to implement the use of Twitter into intervention efforts since permission to use private profile data is not needed. Given that this research indicates that tweeting more frequently about alcohol (and using certain terms) is related to actual drinking cognitions and behaviors, targeting specific key words should reach those most at risk for risky alcohol use.

6. Future directions and limitations

One limitation of the present study is that the collection of tweets were not necessarily lined up temporally with the survey, meaning that the snapshot of tweets obtained may have been posted both before and after the one-time survey was completed by participants. Therefore, while we were able to find associations, longitudinal and daily studies are needed in order to determine the predictive relationships between Twitter activity and high-risk alcohol cognitions, alcohol use, and consequences among adolescents and YAs. In particular, it is important to disentangle the influence of Twitter social networks on adolescent and YA alcohol use. Future research is needed to assess both within- and between-person associations between participants' Twitter activity and drinking cognitions, alcohol use, and consequences as well as the impact of Twitter social networks on these outcomes. Additionally, because the sample in the present study was YAs age 18-20 in the United States, future research should examine whether these same results generalize in both adolescents and those young adults who are at or above the legal drinking age as well as in YAs in other countries. Finally, although Twitter remains an important part of YA's lives, their involvement in other emerging SNS such as Instagram, and Snapchat should be further explored to determine whether the findings from this study are generalizable across SNS platforms or if there is something unique about Twitter that may facilitate the associations between alcohol-related tweets and personal alcohol use.

Acknowledgments

Role of the Funding Source

Data collection and manuscript preparation were supported by National Institute on Alcohol Abuse and Alcoholism Grant R01AA021379. Manuscript preparation was also supported by National Institute on Alcohol Abuse and Alcoholism Grants R00AA020869 and R21AA024163.

References

Atkins DC, Gallop RJ, 2007 Rethinking how family researchers model infrequent outcomes: A tutorial on count regression and zero-inflated models. J. Fam. Psychol. 21, 726. doi: 10.1037/0893-3200.21040726. [PubMed: 18179344]

Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG, World Health Organization (WHO), 2001 AUDIT: The alcohol use disorders identification test: Guidelines for use in primary health care. [WWW Document]. URL http://whqlibdoc.who.int/hq/2001/WHO_MSD_MSB_01.6a.pdf.

Bennett S, 2014 Twitter USA: 48.2 million users now, reaching 20% of population by 2018. [WWW Document]. URL https://www.adweek.com/digital/twitter-usa-users/.

Blackwell DL, Lucas JW, Clarke TC, 2014 Summary health statistics for U.S. adults: National Health Interview Survey, 2012. [WWW Document]. Vital Health Stat. 10, 1–161. URL http://www.cdc.gov/nchs/data/series/sr_10/sr10_260.pdf.

Bohn MJ, Babor TF, Kranzler HR, 1995 The Alcohol Use Disorders Identification Test (AUDIT): Validation of a screening instrument for use in medical settings. J. Stud. Alcohol 56, 423–432. [PubMed: 7674678]

Boyle SC, LaBrie JW, Froidevaux NM, Witkovic YD, 2016 Different digital paths to the keg? How exposure to peers' alcohol-related social media content influences drinking among male and female first-year college students. Addict. Behav. 57, 21–29. doi:10.1016/j.addbeh.2016.01.011. [PubMed: 26835604]

Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Bierut LJ, 2015 "Hey everyone, I'm drunk." An evaluation of drinking-related Twitter chatter. J. Stud. Alcohol Drugs Suppl. 76, 635–643. doi:10.15288/jsad.2015.76.635.

Collins RL, Parks GA, Marlatt GA, 1985 Social determinants of alcohol consumption: The effects of social interaction and model status on the self-administration of alcohol. J. Consult. Clin. Psychol. 53, 189–200. doi:10.1037/0022006X.53.2.189. [PubMed: 3998247]

D'Angelo J, Kerr B, Moreno MA, 2014 Facebook displays as predictors of binge drinking: From the virtual to the visceral. Bull. Sci. Technol. Soc. 34, 159–169. doi: http://dx.doi.org/10.1177/0270467615584044. [PubMed: 26412923]

Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M, 2015 Frequency of social media use. [WWW Document]. URL http://www.pewinternet.org/2015/01/09/frequency-ofsocial-media-use-2/.

Ellison NB, Steinfield C, Lampe C, 2007 The benefits of Facebook 'friends:' Social capital and college students' use of online social network sites. J. Comput. Mediat. Commun. 12, 1143–1168. doi: 10.1111/j.1083-6101.2007.00367.x.

Erevik E, Torsheim T, Vedaa O, Andreassen C, Pallesen S, 2017a Sharing alcohol-related content on social networking sites: Frequency, content, and correlates. J. Stud. Alcohol Drugs Suppl. 78, 608–616. doi:10.15288/jsad.2017.78.608.

Erevik EK, Torsheim T, Andreassen CS, Vedaa Ø, Pallesen S, 2017b Disclosure and exposure of alcohol on social media and later alcohol use: A large-scale longitudinal study. Front. Psychol. 8, 1934. doi:https://doi.org/10.3389/fpsyg.2017.01934. [PubMed: 29163308]

Gerrard M, Gibbons FX, Houlihan AE, Stock ML, Pomery EA, 2008 A dual-process approach to health risk decision making: The prototype willingness model. Dev. Rev. 28, 29–61. doi: 10.1016/j.dr.2007.10.001.

Geusens F, Buellens K, 2016 Strategic self-presentation or authentic communication? Predicting adolescents' alcohol references on social media. J. Stud. Alcohol Drugs 78, 124–133. doi:http://dx.doi.org/10.15288/jsad.2017.78.124.

Geusens F, Beullens K, 2017 The reciprocal associations between sharing alcohol references on social networking sites and binge drinking: A longitudinal study among late adolescents. Comput. Human Behav. 73, 499–506. doi:10.1016/j.chb.2017.03.062.

Geusens F, Beullens K, 2016a The association between social networking sites and alcohol abuse among Belgian adolescents: The role of attitudes and social norms. J. Media Psychol. 1–10. doi.org/10.1027/1864-1105/a000196.

Hingson R, White A, 2014 New research findings since the 2007 Surgeon General's call to action to prevent and reduce underage drinking: A review. J. Stud. Alcohol Drugs Suppl. 75, 158–169. doi: 10.15288/jsad.2014.75.158.

Hoffman EW, Austin EW, Pinkleton BE, Austin BW, 2017 An exploration of the associations of alcohol-related social media use and message interpretation outcomes to problem drinking among college students. J. Health Commun. 32, 864–871. doi:10.1080/10410236.20161195677.

Huang GC, Unger JB, Soto D, Fujimoto K, Pentz MA, Jordan-Marsh M, Valente TW, 2014 Peer influences: The impact of online and offline friendship networks on adolescent smoking and alcohol use. J. Adolesc. Health 54, 508–514. doi:10.1016/j.jadohealth.2013.07.001. [PubMed: 24012065]

Huberman BA, Romero DM, Wu F, 2008 Social networks that matter: Twitter under the microscope. First Monday 14. doi: 10.5210/fm.v14i1.2317.

Krauss MJ, Grucza RA, Bierut LJ, Cavazos-Rehg PA, 2017 "Get drunk. Smoke weed. Have fun": A content analysis of tweets about marijuana and alcohol. Am. J. Health Promot. 31, 200–208. doi: 10.4278/ajhp.150205-QUAL-708. [PubMed: 26559715]

Kwak H, Lee C, Park H, Moon S, 2010 What is Twitter, a social network or a news media? Paper presented at the Proceedings of the 19th international conference on World wide web, Raleigh, NC.

Lee CM, Geisner IM, Lewis MA, Neighbors C, Larimer ME, 2007 Social motives and the interaction between descriptive and injunctive norms in college student drinking. J. Stud. Alcohol Drugs Suppl. 68, 714–721. doi:10.15288/jsad.2007.68.714.

Litt DM, Lewis MA, 2015 Examining the role of abstainer prototype favorability as a mediator of the abstainer-norms–drinking-behavior relationship. Psychol. Addict. Behav. 29, 467–472. doi: 10.1037/adb0000042. [PubMed: 25437152]

Litt DM, Stock ML, 2011 Adolescent alcohol-related risk cognitions: The roles of social norms and social networking sites. Psychol. Addict. Behav. 25, 708–713. doi:10.1037/a0024226. [PubMed: 21644803]

Marengo D, Giannotta F, Settanni M, 2017 Assessing personality using emoji: An exploratory study. Pers. Individ. Dif. 112, 74–78. doi:10.1016/j.paid.2017.02.037.

McCormick T, Lee H, Cesare N, Shojaie Emma A, Spiro S, 2015 Using twitter for demographic and social science research: Tools for data collection. Sociol. Methods Res. 46, 390–421. doi: 10.1177/0049124115605339. [PubMed: 29033471]

McCreanor T, Lyons A, Griffin C, Goodwin I, Barnes HM, Hutton F, 2013 Youth drinking cultures, social networking and alcohol marketing: Implications for public health. Crit. Public Health 23, 110–120. doi:10.1080/09581596.2012.748883

Moreno MA, Christakis DA, Egan KG, Brockman LN, Becker T, 2012 Associations between displayed alcohol references on Facebook and problem drinking among college students. Arch. Pediatr. Adolesc. Med. 166, 157–163. [PubMed: 21969360]

Moreno MA, Arseniev-Koehler A, Litt D, Christakis D, 2016 Evaluating college students' displayed alcohol references on Facebook and Twitter. J. Adolesc. Health 58, 527–532. doi:http://dx.doi.org/10.1016/j.jadohealth.2016.01.005. [PubMed: 26995291]

Moreno MA, Cox ED, Young HN, Haaland W, 2015 Underage college students' alcohol displays on Facebook and real-time alcohol behaviors. J. Adolesc. Health 56, 646–651. doi:10.1016/j.jadohealth.2015.02.020. [PubMed: 26003580]

Moreno MA, Whitehill JM, 2014 Influence of social media on alcohol use in adolescents and young adults. Alcohol Res. 36, 91–100. [PubMed: 26259003]

Nesi J, Rothenberg WA, Hussong AM, Jackson KM, 2017 Friends' alcohol-related social networking site activity predicts escalations in adolescent drinking: Mediation by peer norms. J. Adolesc. Health 60, 641–647. doi:10.1016/j.jadohealth.2017.01.009. [PubMed: 28325545]

O'Malley PM, Johnston LD, 2002 Epidemiology of alcohol and other drug use among American college students. J. Stud. Alcohol Drugs Suppl. 14, 23–39. doi:10.15288/jsas.2002.s14.23.

Smith A, Anderson M, 2018 Social Media Use in 2018. Pew Research Center http://www.pewinternet.org/2018/03/01/social-media-use-in-2018.

Phua J, Jin SV, Kim JJ, 2017 Gratifications of using Facebook, Twitter, Instagram, or Snapchat to follow brands: The moderating effect of social comparison, trust, tie strength, and network homophily on brand identification, brand engagement, brand commitment, and membership intention. Telemat Inform. 34, 412–424. doi: 10.1016/j.tele.2016.06.004.

R Core Team., 2013 R: A language and environment for statistical computing. R.J. Vienna, Austria [WWW Document]. URL http://www.R-project.org/.

Read JP, Kahler CW, Strong DR, Colder CR, 2006 Development and preliminary validation of the young adult alcohol consequences questionnaire. J. Stud. Alcohol 67, 169–177. doi:10.15288/jsa.2006.67.169. [PubMed: 16536141]

Read JP, Wood MD, Davidoff OJ, McLacken J, Campbell JF, 2002 Making the transition from high school to college: The role of alcohol-related social influence factors in students' drinking. J. Subst. Abuse 23, 53–65. doi:10.1080/08897070209511474.

Steele JR, 2005 Children, teens, families, and mass media: The millennial generation [Review of the book Children, teens, families, and mass media: The millennial generation]. Journal Mass. Commun. Q. 82, 199–203.

Stoddard SA, Bauermeister JA, Gordon-Messer D, Johns M, Zimmerman MA, 2012 Permissive norms and young adults' alcohol and marijuana use: The role of online communities. J. Stud. Alcohol Drugs Suppl. 73, 968–975. doi:10.15288/jsad.2012.73.968.

Substance Abuse and Mental Health Services Administration (SAMHSA), 2014 Results from the 2013 National Survey on Drug Use and Health: Summary of national findings. [WWW Document]. URL http://www.samhsa.gov/data/sites/default/files/NSDUHresultsPDFWHTML2013/Web/NSDUHresults2013.pdf.

Thompson CM, Romo LK, 2016 College students' drinking and posting about alcohol: Forwarding a model of motivations, behaviors, and consequences. J. Health Commun 21, 688–695. doi:http://dx.doi.org/10.1080/10810730.2016.1153763. [PubMed: 27186824]

Tucker JS, Miles JN, and D'Amico EJ, 2013 Cross-lagged associations between substance use-related media exposure and alcohol use during middle school. J. Adolesc. Health 53, 460–464. doi: 10.1016/j.jadohealth.2013.05.005. [PubMed: 23770074]

United States Department of Health and Human Services (USDHHS), 2015 Healthy People 2020 Topics and Objectives. [WWW Document]. URL https://www.healthypeople.gov/2020/topics-objectives.

West JH, Hall PC, Hanson CL, Prier K, Giraud-Carrier C, Neeley ES, Barnes MD, 2012 Temporal variability of problem drinking on Twitter. Open J. Prev. Med. 2, 43–38. doi:10.4236/ojpm.2012.21007.

Westgate EC, Neighbors C, Heppner H, Jahn S, Lindgren KP, 2014 “I will take a shot for every ‘like’ I get on this status”: Posting alcohol-related Facebook content is linked to drinking outcomes. J. Stud. Alcohol Drugs 75, 390–398. doi:http://dx.doi.org/10.15288/jsad.2014.75.390. [PubMed: 24766750]

White AM, Macinnes E, Hingson RW, Pan IJ, 2013 Hospitalizations for suicide-related drug poisonings and co-occurring alcohol overdoses in adolescents (ages 12–17) and young adults (ages 18–24) in the United States, 1999–2008: Results from the nationwide inpatient sample. Suicide Life Threat. Behav 43, 198–212. doi:10.1111/sltb.12008. [PubMed: 23356834]

Highlights

- Alcohol-related tweets, including alcohol emojis, are common among young adults.
- Proportion of alcohol-related twitter posts are associated with drinking outcomes.
- Implications for how Twitter data can be used for surveillance and intervention.

Table 1:

Demographic composition of all participant and participant subgroups

	All	Participants Active on Twitter	Participants Not Active on Twitter	p-value	Participants for Whom Twitter Data Was Collected	Participants for Whom Twitter Data Was Not Collected	p-value
N	1058	406	652	-	186	872	-
Age (years +/− SD)	19.18 (0.79)	19.19 (0.81)	19.18 (0.78)	0.804	19.17 (0.82)	19.19 (0.79)	0.850
Female at Birth	52.6%	55.9%	50.4%	0.309	54.3%	52.2%	0.795
Male at Birth	46.1%	43.6%	47.7%		45.7%	46.2%	
Asian	14.7%	14.5%	14.8%	0.150	12.9%	15.1%	0.533
African American	19.6%	20.9%	18.7%		21.0%	19.3%	
Caucasian	48.1%	51.0%	46.3%		48.9%	47.9%	
American Indian or Alaskan Native	1.3%	0.5%	1.8%		0.5%	1.5%	
Native Hawaiian or Pacific Islander	0.9%	0.2%	1.2%		0%	1.0%	
More Than One Race	9.4%	7.9%	10.3%		10.2%	9.2%	
Other	4.8%	4.9%	4.8%		6.5%	4.5%	
Other/Mixed (above two rows summed)	14.2%	12.8%	15.0%		16.7%	13.6%	
Hispanic/Latino	16.4%	14.5%	17.5%	0.188	14.5%	16.7%	0.476
Not Hispanic	81.9%	85.0%	80.1%		84.9%	81.3%	
Not Student	19.2%	16.5%	20.9%	0.074	13.4%	20.4%	0.026*
In High School	0.9%	1.5%	0.6%	0.098	2.2%	0.7%	0.490
Pursuing GED	0.2%	0%	0.3%		0%	0.2%	
In Community College	17.9%	15.8%	19.2%		18.8%	17.7%	
In Vocational/Tech School	2.4%	2.2%	2.5%		2.2%	2.4%	
In 4-year University/College	54.0%	59.1%	50.8%		60.2%	52.6%	
In Graduate/Professional School	0.5%	0%	0.8%		0%	0.6%	
Other	0.6%	0.7%	0.5%		1.1%	0.5%	

Table 2:

Twitter posts (i.e., tweets) were coded for alcohol-related terms and emojis.

Proportions, numbers, and users of alcohol-related terms/emojis in tweets are listed below in the following tabulation:

Term	Number of Tweets	Percentage of Tweets (%)	Unique Users	Proportion of Users (%)
drunk	790	0.234	98	52.69
drink	671	0.199	117	62.90
drinking	356	0.105	92	49.46
beer	280	0.083	72	38.71
bar	223	0.066	85	45.70
shots	214	0.063	80	43.01
alcohol	185	0.055	58	31.18
wine	169	0.050	51	27.42
vodka	111	0.033	49	26.34
wasted	107	0.032	63	33.87
liquor	74	0.022	33	17.74
hangover	65	0.019	36	19.35
hungover	46	0.014	25	13.44
tequila	43	0.013	25	13.44
whiskey	43	0.013	12	6.45
pregame	30	0.009	22	11.83
tipsy	26	0.008	20	10.75
booze	20	0.006	14	7.53
cocktail	12	0.003	10	5.38
Emoji	**Number of Tweets**	**Percentage of Tweets (%)**	**Unique Users**	**Proportion of Users (%)**
clinking beer mugs	145	0.043	39	20.97
beer mug	39	0.012	26	13.98
wine glass	39	0.012	21	11.29
tropical drink	37	0.011	25	13.44
cocktail glass	21	0.006	16	8.60

Table 3:

OLS Regression outcomes for willingness to drink as an outcome are listed below in the following tabulation. Model 1 does not include tweeting behavior as an independent variable. Coefficients are standardized. (Reference dummies include: Gender – *Female*, Race – *Asian*, Ethnic – *Not Hispanic/Latino*):

Regression Criterion	Predictor	Model 1		Model 2	
		β	p > |t|	β	p > |t|
Willingness	Intercept	-	0.548	-	0.595
	Age	0.125	0.136	0.108	0.171
	Gender – *Male*	−0.0369	0.667	0.022	0.786
	Gender – *Female to Male*	−0.179	0.034*	−0.141	0.076
	Race – *Black/African American*	−0.253	0.040*	−0.264	0.023*
	Race – *Caucasian*	0.076	0.561	−0.023	0.885
	Race – *More than One*	0.051	0.637	−0.152	0.141
	Race – *Other*	0.111	0.307	0.096	0.346
	Ethnic – *Hispanic/Latino*	−0.100	0.285	−0.095	0.280
	Proportion of Alcohol Tweets	-	-	0.352	0.0**

Table 4:

OLS Regression outcomes for drinks per week, outcomes, and problem drinking as outcomes are listed below in the following tabulation. Model 1 does not include tweeting behavior as an independent variable. Coefficients are not standardized. (Reference dummies include: Gender – *Female*, Race – *Asian*, Ethnic – *Not Hispanic/Latino*):

Regression Criterion	Predictor	Model 1		Model 2	
		B	p > |z|	B	p > |z|
Drinks per week	Intercept	−4.782	0.178	−2.038	0.565
	α	2.912	0.0**	2.768	0.0**
	Age	0.301	0.098	0.150	0.409
	Gender – *Male*	0.238	0.377	0.374	0.165
	Gender – *Female to Male*	−0.806	0.544	0.090	0.950
	Race – *Black/African American*	0.549	0.249	0.451	0.334
	Race – *Caucasian*	0.919	0.029*	0.568	0.184
	Race – *American Indian/Alaska Native*	0.991	0.575	0.889	0.624
	Race – *More than one*	0.491	0.379	0.359	0.514
	Race – *Other*	1.301	0.095	1.142	0.128
	Ethnic – *Hispanic/Latino*	0.012	0.979	0.099	0.825
	Proportion of Alcohol Tweets	-	-	0.252	0.033*
Consequences	Intercept	−0.306	0.894	−0.314	0.889
	α	1.235	0.0**	1.180	0.0**
	Age	0.092	0.431	0.087	0.452
	Gender – *Male*	−0.238	0.193	−0.169	0.353
	Gender – *Female to Male*	−1.273	0.187	−1.059	0.267
	Race – *Black/African American*	0.200	0.536	0.156	0.623
	Race – *Caucasian*	0.365	0.205	0.213	0.462
	Race – *American Indian/Alaska Native*	0.647	0.591	−0.448	0.701
	Race – *More than one*	0.107	0.781	0.007	0.986
	Race – *Other*	0.640	0.221	0.540	0.291
	Ethnic – *Hispanic/Latino*	−0.161	0.605	−0.079	0.798
	Proportion of Alcohol Tweets	-	-	0.151	0.032*

Regression Criterion	Predictor	Model 1		Model 2	
		B	**p > |z|**	**B**	**p > |z|**
Problem Drinking	Constant	−0.736	0.762	−1.325	0.577
	α	0.723	0.0**	0.654	0.0**
	Age	0.147	0.233	0.162	0.178
	Gender – *Male*	−0.002	0.991	0.167	0.390
	Gender – *FTM*	−1.360	0.084	−1.127	0.143
	Race – *Black/African American*	−1.102	0.003**	−1.028	0.004**
	Race – *Caucasian*	−0.028	0.929	−0.114	0.706
	Race – *American Indian/Alaska Native*	−0.144	0.880	−0.396	0.662
	Race – *More than one*	−0.463	0.252	−0.455	0.239
	Race – *Other*	0.484	0.370	0.397	0.442
	Ethnic – *Hispanic/Latino*	−0.564	0.115	−0.476	0.170
	Proportion of Alcohol Tweets	-	-	0.221	0.009**