

Temporal variability of problem drinking on Twitter

Joshua Heber West^{1*}, Parley Cougar Hall¹, Kyle Prier¹, Carl Lee Hanson¹,
Christophe Giraud-Carrier², E. Shannon Neeley³, Michael Dean Barnes¹

¹Department of Health Science, Brigham Young University, Provo, USA; *Corresponding Author: josh.west@byu.edu

²Department of Computer Science, Brigham Young University, Provo, USA

³Department of Statistics, Brigham Young University, Provo, USA

Received 16 October 2011; revised 24 November 2011; accepted 15 December 2011

ABSTRACT

Twitter is a micro-blogging application, which is commonly used as a way for individuals to maintain social connections. Social scientists have also begun using Twitter as a data source for understanding more about human interactions. There is very little research about Twitter's utility for monitoring health related attitudes, beliefs and behaviors. The purpose of this study was to examine the extent to which individuals tweeted about problem drinking, and to identify if such tweets corresponded with time periods when problem drinking was likely to occur. Data from this study came from tweets originating in one of 9 randomly selected states, one from each of the nine census geographies in the US. Twitter's API was used to collect tweets during the month of October 2010, and again during the time period surrounding New Year's Eve 2010. Keywords were selected which indicated problem-drinking behaviors, and tweets were coded for the presence or absence of these keywords. Twitter users were most likely to tweet about problem drinking on Friday, Saturday and Sunday during the hours from 10 pm to 2 am. Tweets originating during the New Year's Eve holiday (0.53%) were twice as common when compared to tweets during weekends in the month of October (0.34%). Twitter may be a valid data source for social scientists, given that tweets about problem drinking corresponded with expected time periods of actual problem drinking. Furthermore, tweets that mention problem drinking may be problematic for public health if they establish incorrect normative beliefs that such behaviors are acceptable and expected. Social norms interventions may be an effective tool in correcting misperceptions related to problem drinking by informing Twitter followers that problem drinking behaviors are not normative.

Keywords: Twitter; Problem Drinking; Temporal Variance

1. INTRODUCTION

Recent research has called for better integration of theoretical principles to understand online behaviors [1]. Twitter is a social media application where users communicate in brief text posts of 140 characters or less [2], which has become increasingly popular [3]. The Theory of Planned Behavior (TPB) is a theoretical framework that contextualizes the relative importance of Twitter content to a behavior such as problem drinking [4,5].

Within the TPB, a normative belief refers to an individual's perception of a particular behavior, which is influenced by the judgment of significant others (e.g., friends) [5]. Similarly, the subjective norm is an individual's perception of social normative pressures, or relevant others' beliefs that such behavior should or should not be engaged in. The influence on behavior of perceived subjective norms depends upon the importance one attributes to each relevant other's opinion and one's own motivation to comply. Thus the TPB theorizes that an individual's behavior (e.g., binge drinking) is greatly impacted by one's perception of others, especially those considered important. Traditional social networks (e.g., friends, family, peer groups), together with emerging social networks (e.g., chat rooms, social network sites, and Twitter) both contribute to one's perception of normative behavior and the subsequent formation of normative beliefs. Tweepsters can be considered important referents by which followers are making judgments about the tweeted behavior and gauging their motivation to comply, especially given that many Twitter users are following to maintain social connections [2,6]. As the prevalence of tweets about binge drinking increases, followers are more likely to consider the behavior as normative [7]. This type of informal or opportunistic communication occurring on Twitter has the potential to shape normative attitudes and beliefs related to health behaviors in powerful ways [8].

Research has demonstrated temporal variability with

respect to alcohol consumption. Del boca *et al.* [9] showed that weekends are the most common days of the week for alcohol consumption and that rates of binge drinking almost double over the New Year's holiday. Mäkelä *et al.* [10] also observed the weekend trend related to consumption and further demonstrated that alcohol-related deaths increased on major drinking holidays. Lemmens and Knibbe [11] observed a 70 percent increase in alcohol consumption during the final two weeks of December, a time period which is marked by two major holidays. These studies about the epidemiology of temporal variation of alcohol consumption and problem drinking demonstrate a well-established and generally acknowledged social norm of such behaviors during weekends and holidays. To date, however, no research has explored the extent to which microblogging tools such as Twitter may work to reinforce alcohol-related social norms on weekends and holidays.

Recently social scientists have begun to use Twitter as a research tool in examining various social occurrences. Preliminary studies are necessary to establish the validity of Twitter as a research tool, by demonstrating that discussion on Twitter coincides with traditional communication. For example, Golder and Macy [12] reported that mood fluctuations including temporal and seasonal variations broadcast via tweets were consistent with expected patterns. Notably, tweet sentiment was significantly happier on weekends. Whereas these findings are expected and therefore do not extend our knowledge of human behavior per se, they indicate that Twitter may indeed be an emerging tool for monitoring attitudes, beliefs and behaviors. Research has yet to explore Twitter's utility for monitoring or measuring the health related behavior of its users. The current study represents a preliminary attempt to understand the temporal variability of tweets related to problem drinking. Two primary hypotheses guided this study. First, it was predicted that tweets related to problem drinking would significantly increase during nighttime hours and on weekends. Second, it was predicted that tweets related to problem drinking would significantly increase during the New Year's Eve holiday.

2. METHOD

2.1. Sample

Data for this study came from 5,697,008 tweets generated from Twitter users in nine states, randomly selected from each of the federal census divisions in the US. Tweets from Georgia, Idaho, Indiana, Kansas, Louisiana, Massachusetts, Mississippi, Oregon, and Pennsylvania comprised the study sample.

Tweets are publicly available and easily accessed through Twitter's Application Programming Interface (API). An API is an interface, or a set of rules, that allow one program to communicate with and access the resources of another.

For example Twitter's API allows other programs to extract the content and origin of selected tweets. Using the Twitter search API, recent tweets for each state were gathered in 2-minute intervals over a 31-day period from October 5, 2010 through November 3, 2010, and over another 5-day period from December 30, 2010 through January 3, 2011. All of the tweets collected during these time periods were imported into a database, where they were prepared for further statistical analysis with SAS software. All non-English tweets were excluded from the study sample.

2.2. Measures

Tweets were identified that contained words reflective of problem drinking (e.g., drunk), not mere mentions of alcohol (e.g., beer). A slang compilation by the Big Book Bunch group of Alcoholics Anonymous (<http://www.sober.org/Drunk.html>) was used to identify commonly used phrases that generally reflect problem drinking. The Online Slang Dictionary (<http://onlineslangdictionary.com>) was also used to identify additional synonyms by querying the database for the word drunk. The resulting terms from these two searches are presented in **Table 1**. The database of tweets was queried to identify the presence or absence of at least one of the terms in each tweet. A value of 1 was assigned to tweets that contained at least one of the terms, while a value of 0 was assigned to tweets that contained none of the terms. Tweets that contained references to problem drinking are hereafter referred to simply as alcohol-related tweets. Note however that, as stated above, they reflect problem drinking and not mere mentions of alcohol or beer, which would have resulted in a coding of 0.

Table 1. Synonyms of drunk.

Drunk
Wasted
Tipsy
Intoxicated
Hammered
Sauced
Buzzed
Trashed
Under the influence
Blacked out
Gassed
Zooted
Juiced
Designated driver
Shit-faced
Plastered
Inebriated
Cock eyed
Cockeyed
Shwasted
Zonked

Twitter’s search API geocode parameter was used to ensure that tweets originated within the nine randomly selected states. The geocode parameter requires latitude, longitude, and a radial distance to capture tweets within a circular area

(<http://apiwiki.twitter.com/w/page/22554756/Twitter-Search-API-Method:-search>). Codes were generated to ensure that a minimum of 90% area coverage for each state was achieved. Additionally, the 2000 US Census results were consulted to confirm that each state’s top ten most populous cities were included in the geocode regions. All of this information was used to ascertain longitude and latitude, or the exact location of origin for each tweet. Information related to the time of the tweet, day of the week, and the actual text of the tweet, automatically populated the specialized database.

2.3. Statistical Analysis

In testing both hypotheses, chi-square test statistics were computed to test the differences in the percentage of al-

cohol-related tweets, which were compared by time of day, day of the week, and between the New Year’s holiday and weekend tweets from October 2010. All tweet times were adjusted to reflect the local time of the time zone in which the tweet was published.

3. RESULTS

Figures 1 and 2 address this study’s first hypothesis related to the percentage of alcohol related tweets during nighttime hours and weekends. **Figure 1** displays tweets by time of day for the time period from October 5, 2010 to November 3, 2010. Alcohol-related tweets were most common between the hours of 9 PM and 2 AM ($p < 0.0001$). **Figure 2** shows that Twitter users were most likely to tweet alcohol-related content on Friday, Saturday or Sunday ($p < 0.0001$).

Figures 3 and 4 address this study’s second hypothesis related to the comparison of alcohol related tweets during the New Year’s holiday. The appropriateness of such a comparison was aided in that 2010-2011 New Year’s holiday

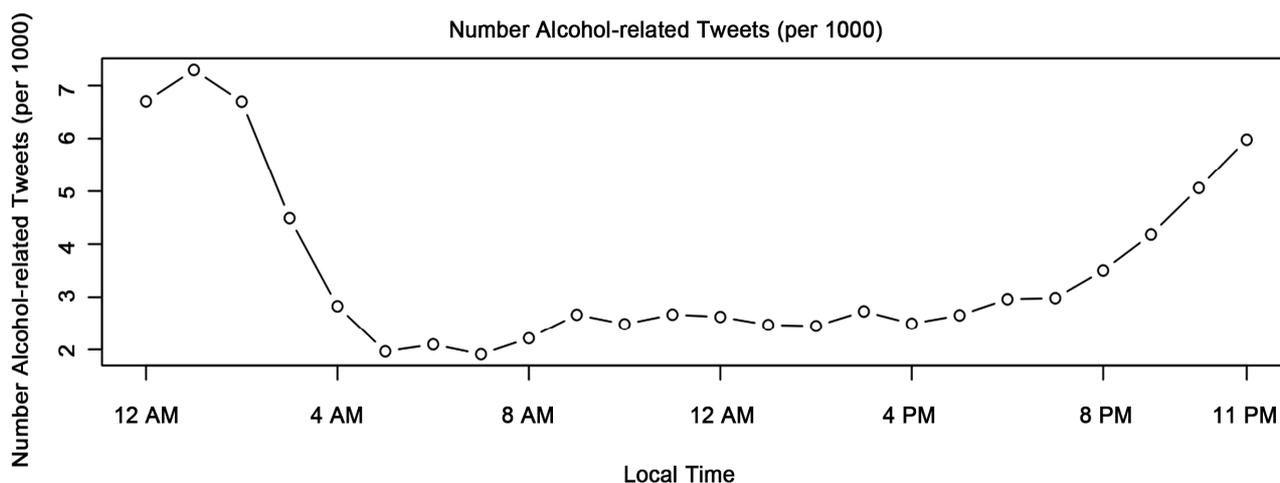


Figure 1. Alcohol-related tweets by time of day, October 5, 2010 to November 3, 2010.

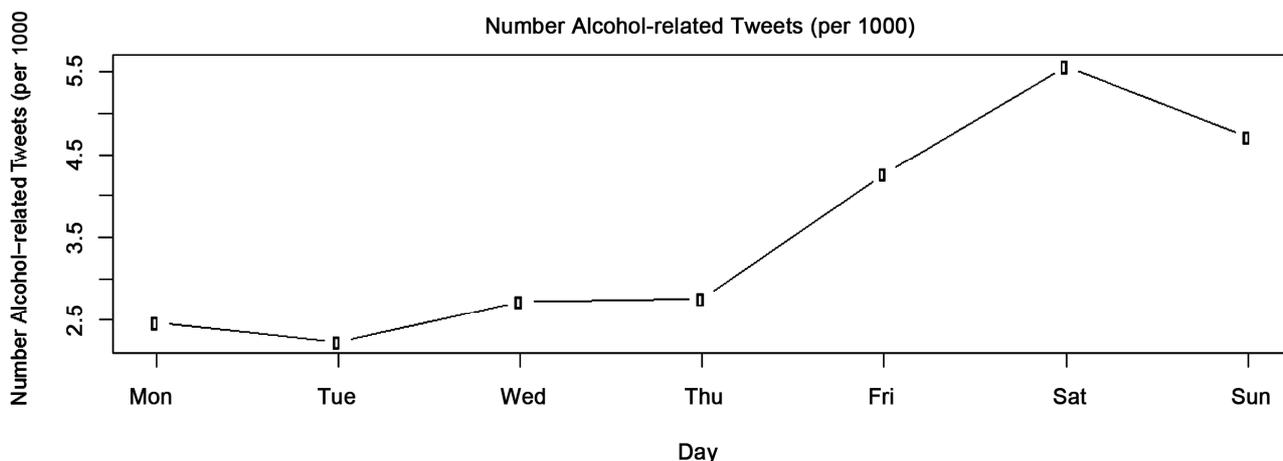


Figure 2. Alcohol-related tweets by day of the week, October 5, 2010 to November 3, 2010.

fell on a weekend. There were a total of 4,727,046 tweets from October 5, 2010 to November 3, 2010, 16,046 of which were alcohol-related (0.34%). From December 30, 2010 to January 3, 2011 there were 969,962 tweets and 5132 (0.53%) were alcohol-related. **Figure 3** shows that on New Year's Eve (Friday) and New Year's Day (Saturday), alcohol-related tweets were significantly higher than on similar weekend days during the month of October ($p < 0.0001$). The timing of the alcohol-related tweets in relation to the time of day is shown in **Figure 4**. As the New Year is introduced (Sat 12 AM, January 1st, 2011), alcohol-related tweets are at their highest level.

4. DISCUSSION

The results from this study confirm both study hypotheses and indicate that during historical drinking time periods (e.g., nights, weekends, and New Year's) alcohol related tweets were most common. Whereas these findings

do not extend our understanding of human behavior per se, this sort of validity testing of an emerging data source provides preliminary confirmation of its potential value in monitoring and understanding health behaviors. Had the findings from this study been inconsistent with the study hypotheses, concerns would have been raised about the utility of microblogging and social media content as a surveillance tool for social scientists. Other recent validity tests of social media sites have further confirmed that individuals reporting about alcoholism [13] and depression [14,15] on Facebook, indeed, suffer from those conditions.

The objective of the study being presented here is consistent with the recent call by Neighbors *et al.* [20] for research that evaluates normative influences on drinking. While tweets simply mirror expected alcohol related patterns, they do so on a grand scale. Whereas social norms and expectations have been communicated at a localized level in the past (e.g., face to face, telephone, social gathering), social media platforms like Twitter now provide users the ability to instantly communicate on a global level. This ability to communicate one's alcohol related behavior to large groups of people has the ability to further normalize such behaviors. For this reason, alcohol-related Twitter communications might be of concern to individuals dedicated to correcting the misperceptions of alcohol-related social norms, which are associated with increased problem drinking [16]. Beliefs related to others' alcohol consumption are strongly related to alcohol use. Individuals who overestimate drinking norms often drink at higher levels than those who estimate norms more accurately [21]. Normative beliefs are powerful determinants of behavior and serve as a standard against which to compare one's own behavior. In their study of college drinkers, Borsari and Carey [21] concluded that "the more the student perceives others as drinking heavily, or approving of heavy use, the higher personal consumption will be" (p. 402). Normative beliefs supportive of alcohol-related behaviors and exaggerated perceptions of actual peer drinking are serious obstacles in curbing problem drinking and subsequent alcohol-related

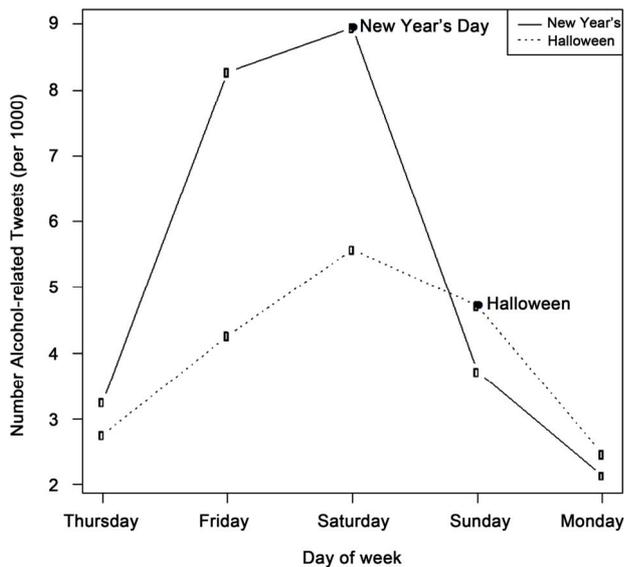


Figure 3. Alcohol-related tweets from weekends in October, 2010 and New Year's Eve, 2010.

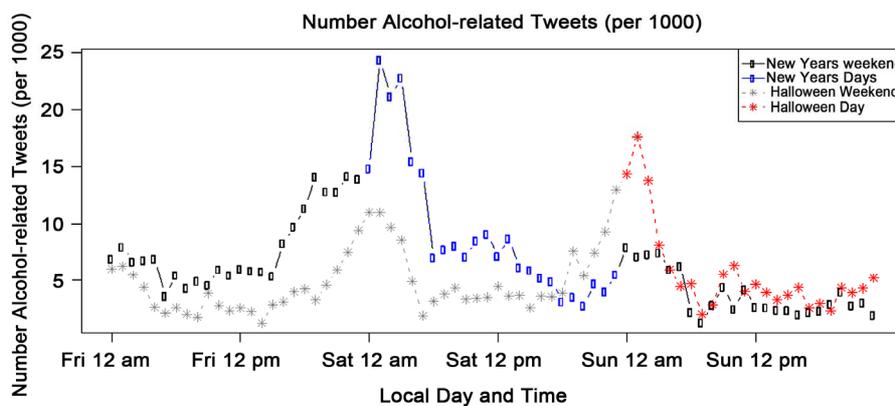


Figure 4. Alcohol-related tweets by time of day from weekends in October, 2010 and New Year's Eve, 2010.

problems. Normative influences may be particularly persuasive to adolescent populations. Recent research has demonstrated that adolescents that use social networking sites are more likely to use tobacco, alcohol and marijuana and 40% report having seen a picture of kids drunk, passed out, or using drugs online [22]. Repeated exposures to such influences may further perpetuate social norms among this age group.

Whereas the overall percentage of alcohol-related tweets in this study was small, the patterns of alcohol-related tweets were observed to be in the expected direction. Furthermore, the actual percentage of individuals that engage in problem drinking is low. Nationally, only 5% of adults report being heavy drinkers and just 15% report binge drinking [17]. In reality it is difficult to speculate on the extent to which Tweepsters may misreport about their problem drinking. Studies of this nature have not yet been conducted, but could be the focus of subsequent research. On the other hand, while it is challenging to know if problem drinking is misreported, twitter research may help avoid certain challenges commonly faced by social scientists such as social desirability response bias, participation bias, and others. Additionally, future studies should build upon recent work by Paul and Dredze [18] and Prier *et al.* [19] to further explore the frequency of prominent health related topics being discussed on Twitter.

The findings from this study should be interpreted in the context of several limitations. Whereas the list of synonyms for problem drinking was very inclusive of both common and uncommon terms, it was not possible to determine in this study if the terms were used in a positive or negative context. However, a post-hoc review of a sample of the Twitter content revealed that most of the tweets referencing problem drinking did so in a negative context (e.g., "I was so drunk I'm sick"). Related to this, it was not possible to determine if tweets were referencing independent or unique problem drinking episodes. For example, multiple users may tweet about news reports of a single celebrity's problem with drinking. Absent more refined metrics, such a scenario could limit Twitter's utility as a tool for monitoring reports of problem drinking. Future studies might benefit from more extensive analyses of the content of alcohol-related tweets, perhaps using sentiment analysis as suggested by Savage [8]. Also, while a comprehensive list of terms was generated to identify alcohol-related tweets, because of the evolving nature of slang words it is likely that not every term used to identify problem drinking was discovered. In addition, as is common in social science studies, there are concerns in the current study related to generalizability of the study findings. It is noted that the current study's findings are a reflection of tweets from individuals that use Twitter, which represent a small, yet rapidly expanding segment of the population.

REFERENCES

- [1] Webb, T., Joseph, J., Yardley, L. and Michie, S. (2010) Using the Internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research*, **12**, e4. [doi:10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)
- [2] Java, A., Song, X., Finin, T. and Tseng, B. (2007) Why we Twitter: Understanding microblogging usage and communities. *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, San Jose, 56-65.
- [3] Smith, C. (2011) Twitter user statistics show stunning growth. *The Huffington Post*. http://www.huffingtonpost.com/2011/03/14/twitter-user-statistics_n_835581.html.
- [4] Ajzen, I. and Fishbein, M. (1975) Belief, attitude, intention and behavior: An introduction to theory and research. Addison-Wesley, Reading.
- [5] Montano, D.E. and Kasprzyk, D. (2008) Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In: Glanz, K., Rimer, B.K. and Viswanath, K., Eds., *Health Behavior and Health Education*, Jossey-Bass, San Francisco, 67-96.
- [6] Krishnamurthy, B., Gill, P. and Arlitt, M. (2008) A few chirps about Twitter. *Proceedings of the 1st Workshop on Online Social Networks*, Seattle, 19-24. [doi:10.1145/1397735.1397741](https://doi.org/10.1145/1397735.1397741)
- [7] Zhao, D. and Rosson, M.B. (2009) How and why people Twitter: The role that micro-blogging plays in informal communication at work. *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, Sanibel Island, 243-252.
- [8] Savage, N. (2011) Twitter as medium and message. *Communications of the ACM*, **54**, 18-20.
- [9] Del Boca, F.K., Darkes, J., Greenbaum, P.E. and Goldman, M.S. (2004) Up close and personal: Temporal variability in the drinking of individual college students during their first year. *Journal of Consulting and Clinical Psychology*, **72**, 155-164. [doi:10.1037/0022-006X.72.2.155](https://doi.org/10.1037/0022-006X.72.2.155)
- [10] Mèkelè, P., Martikainen, P. and Nihtilè, E. (2005) Temporal variation in deaths related to alcohol intoxication and drinking. *International Journal of Epidemiology*, **34**, 765-771. [doi:10.1093/ije/dyi025](https://doi.org/10.1093/ije/dyi025)
- [11] Lemmens, P.H.H.M. and Knibbe, R. (1993) Seasonal variation in survey and sales estimates of alcohol consumption. *Journal of Studies on Alcohol*, **54**, 157-163.
- [12] Golder, S.A. and Macy, M.W. (2011) Diurnal and seasonal mood vary with work, sleep, and day length across diverse cultures. *Science*, **333**, 1878-1881.
- [13] Moreno, M.A., Christakis, D.A., Egan, K.G., Brockman, L.N. and Becker, T. (2011) Associations between displayed alcohol references on Facebook and problem drinking among college students. *Archives of Pediatric and Adolescent Medicine*, E1-E7.

- [14] Moreno, M.A., Christakis, D.A., Egan, K.G., Jelenchick, L.A., Cox, E., Young, H., Villiard, H. and Becker, T. (2011) A pilot evaluation of associations between displayed depression references on Facebook and self-reported depression using a clinical scale. *Journal of Behavioral Health Services & Research*, 1-9. [doi:10.1007/s11414-011-9258-7](https://doi.org/10.1007/s11414-011-9258-7)
- [15] Moreno, M.A., Jelenchick, L.A., Egan, K.G., Cox, E., Young, H., Gannon, K.E. and Becker, T. (2011) Feeling bad on Facebook: Depression disclosures by college students on a social networking site. *Depress and Anxiety*, **28**, 447-455. [doi:10.1002/da.20805](https://doi.org/10.1002/da.20805)
- [16] Jung, T., Shim, W. and Mantaro, T. (2010) Psychological reactance and effects of social norms messages among binge drinking college students. *Journal of Alcohol and Drug Education*, **54**, 7-18.
- [17] Centers for Disease Control and Prevention (2010) Alcohol and public health. <http://www.cdc.gov/alcohol/>
- [18] Paul, M.J. and Dredze, M. (2011) You are what you tweet: Analyzing Twitter for public health. *Proceedings of the Fifth International Conference on Weblogs and Social Media*, Barcelona, 265-272.
- [19] Prier, K.W., Smith, M.S., Giraud-Carrier, C. and Hanson, C.L. (2011) Identifying health-related topics in Twitter: An exploration of tobacco-related tweets as a test topic. *Proceedings of the 4th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction*, College Park, 18-25.
- [20] Neighbors, C., Dillard, A.J., Lewis, M.A., Bergstrom, R.L. and Neil, T.A. (2006) Normative misperceptions and temporal precedence of perceived norms and drinking. *Journal of Studies on Alcohol*, **67**, 290-299.
- [21] Borsari, B. and Carey, K. (2003) Descriptive and injunctive norms in college drinking: A meta-analytic integration. *Journal of Studies on Alcohol*, **64**, 331-341.
- [22] The National Center on Addiction and Substance Abuse at Columbia University (2011) National Survey of American attitudes on substance abuse XVI: Teens and parents. <http://www.casacolumbia.org/templates/NewsRoom.aspx?articleid=648&zoneid=51>
- [23] Berkowitz, A.D. (2004) An overview of the social norms approach. In: Lederman, L. and Stewart, L., Eds., *Changing the Culture of College Drinking*, Hampton Press, Cresskill, 193-214.
- [24] Joinson, A.N. (2001) Self disclosure in computer mediated communication: The role of self awareness and visual anonymity. *European Journal of Social Psychology*, **31**, 177-192. [doi:10.1002/ejsp.36](https://doi.org/10.1002/ejsp.36)