Is Yik Yak A Platform for Political Communication? Exploring College Students' Communication on an Emergent Social Media Platform

Abstract

This study used big data and computational methods to explore the degree to which college students use Yik Yak for political expression. While the number of political messages was small relative to the entire corpus, the data suggested that political communication on Yik Yak is responsive to major political events. Moreover, the data suggested that the percentage of political commentary on Yik Yak is negatively associated with size of the undergraduate population, the percent of classes with greater than 50 students, and the number of Greek organizations on campus.

Keywords: Political Communication, Yik Yak, Social Media, College Students, Social Capital, Big Data, Computational Social Science

Is Yik Yak A Platform for Political Communication? Exploring College Students' Communication on an Emergent Social Media Platform

Social media platforms such as Facebook and Twitter have become an object of exploration for researchers interested in political discussion. This scholarly emphasis on social media has been driven, in large part, by the rapid adoption of social networking technologies, particularly among young adults. Available research suggests that young adults use social media for a wide array of political behaviors related to both information seeking and self-expression. Understanding the relationship between young adults' media use and political behaviors is especially important because research suggests that the habits underlying lifelong political engagement develop during one's college years (Chaffee, Ward, & Tipton, 1981; Diddi & LaRose, 2006).

The present study used computational social science techniques to explore the democratic potential of Yik Yak, an emergent social media platform that is especially popular among college students. Yik Yak is a location-based messaging service that allows users to anonymously post content (Northcut, 2015). Acquiring a better understanding of the degree to which Yik Yak is used for political purposes is important as it continues to increase in popularity.

In this study, we explored three specific research questions. First, we set out to better understand the degree to which college students generate politically oriented content on Yik Yak. We also assessed the degree to which responsiveness varies during nationwide political events. We also investigated the degree to which campus-wide contextual factors related to university profile, university learning environment, and social context could be used to describe political behavior on Yik Yak. To answer these research questions, we collected a month's worth of Yik Yak content from 63 major universities located in the United States (over 1.1 million individual messages). This data was collected in January of 2015, which allowed us to assess the degree to

which a national political event (President Obama's State of the Union address) would influence political discussion on Yik Yak. Using the collected data, we employed computer-assisted content analysis techniques to explore the degree to which political expression is present on Yik Yak. Finally, the work investigated the relationship between contextual factors (i.e., college-level data points) and political expression on Yik Yak.

Literature Review

Yik Yak

Yik Yak is a spatially bounded, anonymous messaging application that is primarily used by college students (Huet, 2014; Smith, 2015). The mobile phone-based application allows users to anonymously post messages, or "Yaks." Although users can post pictures, the majority of content posted on Yik Yak is text-based and limited to 200 characters. Depending upon site traffic, users must be within a radius of 1.5 to 10 miles of a central location (i.e., a college campus). Currently, Yik Yak is the most frequently downloaded anonymous social app in Apple's App Store, which reflects its enormous popularity among college students (Mahler, 2015). Users can "upvote" messages as a positive response and "downvote" yaks messages as a negative response (Northcut, 2015). Messages receiving an aggregate of five downvotes disappear from view. Otherwise, user messages stay visible from several hours to several weeks, depending upon the volume of posted content.

College students can use Yik Yak to discuss upcoming finals, ask a question about campus life, or make a joke about a rival school (Mahler, 2015). Notably, serious concerns have emerged over the application's proclivity to harbor defamatory, uncivil, homophobic, and misogynistic discussion. Yik Yak has also been used to anonymously post threatening content. According to Mahler (2015), Yik Yak has been used to issue violent threats on more than a

dozen college campuses. In one case, a Yik Yak user suggested that fellow users partake in a gang rape at the school's women's center. Concern over the nature of student conversation has led to organizations such as the Human Rights Campaign and the National Organization for Women to ask the Department of Education to address harassment and other verbally abusive behaviors on Yik Yak (Hess, 2015; Conner, 2015). In light of such behaviors, scholars, commentators, and community activists have primarily focused on the negative social effects associated with Yik Yak. However, as college students continue to adopt and use Yik Yak, it is important also appraise the degree to which application can facilitate potentially positive social outcomes. One such outcome, contribution to the political public sphere, is the focus of the current study.

Political Expression and Social Media

As it specifically relates to political behaviors, prior research has linked factors such as community social status, community size, and neighborhood structure to both on and offline political involvement (e.g., Boardman & Robert, 2000; Glascock, 2014; Kwak, Williams, Wang, & Lee, 2005). Vargo and Hopp (2015), for instance, used a big data approach to understand the relationship between community factors and political incivility on Twitter. The results suggested aggregated levels of education exerted a negative effect on political incivility – in other words, residents in better educated communities were more likely to discuss politics using civil language than those less educated communities.

Although scholars have yet to assess the degree to which Yik Yak is used for political communication, prior research has explored both why and how citizens use Internet-based media platforms for political purposes. Research in this area tends to conceptualize political participation in terms of both content acquisition and content creation (e.g., Gil de Zúñiga,

Molyneux, & Zheng, 2014; Shah, Cho, Eveland, & Kwak, 2005). As it relates to political participation, people use social media for a variety of activities, including sharing opinions about current events, commenting on public officials, and collectively organizing in response to proposed policies and initiatives (Gil de Zúñiga, Molyneux, & Zheng, 2014). In this study, we focus specifically on political expression, which in the current context can be defined as the posting of political commentary by individual site users.

Given the widespread adoption of social media, investigation of the discursive democratic potential of these platforms has been conducted. On one hand, scholars note that the accessible nature of the Internet has the potential widen the public sphere. This would allow for the expression of additional perspectives that have traditionally absent from institutional news sources (Yamamoto, Kushin, & Dalisay, 2015; Correa & Jeong, 2011). Moreover, online political conversation may be less resource intensive than offline political conversation (Valenzuela, Kim, Gil de Zúñiga, 2012), resulting in wider participation. Conversation is important because research suggests that people who talk about politics are more likely to engage in political activities (Shah, Cho, Eveland, & Kwak, 2005; Valenzuela, Kim, Gil de Zúñiga, 2012). On the other hand, researchers have observed instances where the democratic potential of social media is left unrealized (Papacharissi, 2004). Not all popular online platforms will contain political discussion and be responsive to political events.

One of the primary goals of this study was to examine the degree to which Yik Yak, an emergent social media platform, facilitates college students' political expressions. Therein, we were interested in exploring both the frequency of political discussion on Yik Yak.

RQ1: What percentage of Yik Yak discussion is devoted to political expression?

Moreover, scholars have noticed that the ~~frequency~~ volume of political talk on social media is responsive to major offline social and political events. ~~not equal for all issues and events.~~ In one prior study, Vargo, Shaw and Basilaia (2015) tracked two ongoing ~~issues~~situations. The first, the mortgage and housing crisis, featured relatively infrequent newsworthy developments and thus news coverage was primarily "speculative and debate-driven" (p. 225). ~~was subject to continued discussion and debate.~~ The second scenario, the BP oil spill, featured many newsworthy sub-events ~~contained real-world cues with many major events~~ (i.e., repeated ~~a~~ failed ~~cap~~ attempts to ~~stop~~ cap the spill), thus resulting in news coverage that was fueled by the consistent and continued emergence of cues. ~~. For 90 days on Twitter, cable television and large newspapers were monitored.~~ Generally speaking, the results suggested that the volume of Twitter-based discussion was, indeed, linked to offline occurrences. Therein, the data further suggested that scenarios featuring many real-word cues may be accompanied by volatile levels of social media conversation. Specifically, situations governed by the rapid and continued emergence of newsworthy information were likely to feature stark increases in Twitter conversation volume while scenarios with fewer news cues were characterized by a slow entropy in online conversation volume. ~~The number of messages in response to news coverage in volume and relative percentage were highest when major events occurred. Twitter was more responsive and sparked more conversation about major events, not ongoing issues. This happened regardless of relative coverage in the media.~~ Noting this difference, we were interested in exploring the degree to which general political talk volume on Yik Yak was associated with major political events~~Yik Yak's responsiveness to major events~~, and how those periods might

differ from ~~times in which ongoing discussion about political issues was occurring~~periods that did not feature major offline news cues. ~~.~~

RQ2: Is the amount of political discussion on Yik Yak associated with major offline political events (i.e., the State of the Union Address)? ~~Is political discussion on Yik Yak responsive to national political events?~~ If so, to what degree?

Potential Contextual Correlates of Political Conversation on Yik Yak

In addition to exploring the relative frequency and responsiveness of political expression on Yik Yak, this study also explored the degree to which contextual factors associated with on-campus life were associated with aggregated patterns of political expression. Central to this line of inquiry is Dahlgren's (2000) *civic culture*, which suggests that surrounding social and cultural factors effectively constitute the pre-conditions for democratic participation. In other words, community profile factors fundamentally influence political behavior, including those behaviors related to citizen engagement in the online public sphere. This study focused on three contextual factors associated with campus life: university demographic profile, campus learning environment, and on-campus social factors. These dimensions are reviewed in the following section.

University demographics.

On the individual level, research indicates that social media usage may differ by demographic. For instance, African-American teenagers are more likely than white teenagers to use to Twitter (Madden, et al, 2013) and women have been found to be heavier users than men (Hargittai & Litt, 2011). Research has also shown that age cohorts use social media platforms differentially and with varying frequencies (Duggan et al, 2015). Research suggests that these individual-level differences in social media usage may also be apparent at the community level.

For example, a recent study by Murthy, Gross, and Pensavalle (2015) used big data to explore the relationship between Twitter usage intensity and community demographics. Here, the authors found a relationship between community demographics and tweet interval, concluding that African-American users living in densely populated urban areas may be the most active on Twitter.

Despite the foregoing, relatively little is known about the relationship between community-level characteristics and *political involvement* on social media. Thus, in the current study, we set out to explore if university-level demographic factors are related to political engagement on Yik Yak. Considering the unique characteristics of college campuses across the United States, the following factors were considered: school size, gender breakdown, average age, race share (African-American, Asian, Hispanic, and white), attendance cost, and public or private status. These factors are consistent with prior social media research on demographics, which has, as reviewed above, focused on age, gender, race, and socio-economic status.

University learning environment.

Research has suggested that college attendance has an appreciable impact on political values and involvement (e.g., Hillygus, 2005; McClintock & Turner, 1962). Less is known about the degree to which individual university learning environments affect either political participation generally or political expression specifically. Addressing this question may be important. Prior studies have concluded that the learning environments that surround college students influence student-centered outcomes, including studying habits (Lizzio, Wilson, & Simons, 2002), performance (Kuh, 1993), involvement (Karp & Yoels, 1976), and social integration (Braxton & McClendon, 2001). Thus, one of the goals of this study was to assess a given college's learning environment and its association with online political conversation.

University learning environments can be understood in both institutional/contextual (i.e., classroom size) and perceptive (i.e., student satisfaction) terms. For instance, Lizzio, Wilson, and Simons (2002) used a self-report measure to better understand college students' perceptions of their surrounding learning environment. From an institutional angle, scholars have used measures such as student retention (Braxton & McClendon, 2001), classroom size (Pulvers & Diekhoff, 1999), and graduation rate (Lau, 2003) as indicators of productive learning environments. Considered as a whole, research in this area suggests that college students who have access to ample faculty, administrative, and material resources tend to perform better and demonstrate higher levels of engagement with their studies and campus life. Therefore, in the current study, factors such as freshman retention rate, average class size, and graduation rate were used to explore the relationship between college students' learning environment and the frequency of political communication on Yik Yak.

University social environment.

This study's approach to revealing the relationship between campus social environment and online political communication is heavily influenced by the theory of social capital. Social capital refers to individual citizens' ability to access "resources embedded in their social networks" (Ellison, Vitak, Gray, & Lampe, 2014, p. 856). Social capital is linked to factors such as shared norms, shared language, social trust, and the sense of mutual obligation (Huysman & Wulf, 2004; Ellison, Steinfeld, & Lampe, 2007). It is both created and maintained via social interaction (Ellison, Vitak, Gray, & Lampe, 2014). The ability to form and maintain shared norms and trust is thought to be a pre-requisite for a civil and politically engaged society. According to Newton (2001), social capital helps "sustain civil society and community relations

in a way that generates trust and cooperation between citizens and a high level of civic engagement and participation" (p. 201).

Previous research has explored social capital at both the individual and aggregate levels. On the individual level, social capital has been conceptualized as interpersonal trust (e.g., Lee & Lee, 2010; Shah, 1998), life contentment (e.g., Shah, Kwak, & Holbert, 2001), participation in civic behaviors (e.g., Blanchard & Horan, 1998; Onyx & Bullen, 2000) and involvement with political organizations (e.g., Wellman, Haase, Witte, & Hampton, 2001). On the contextual level, scholars have measured social capital using a variety of indexes, including aggregated composites of individual level variables (Hendryx, Ahern, Lovrich, & McCurdy, 2002), voter turnout (e.g., Chamlin & Cochran, 1995; Rosenfeld, Messner, & Baumer, 2001), racial/ethnic heterogeneity (e.g., Costa & Kahn, 2003; Rupasingha, Goetz, & Freshwater, 2006) and number of civic/social/political groups within a community (e.g., Pretty & Ward, 2001).

To some degree, the relationship between social capital and political participation on social media remains an open question. On one hand, researchers have found links between the accumulation of social capital and use of social media (e.g., Ellison, Steinfeld, & Lampe, 2007; Valenzuela, Park, & Kee, 2009). However, less is known about the degree to which social capital formation, either *online* or *offline*, is associated with social media-based political expression. Therefore, one of the key questions driving the current study is related to the degree to which campus-wide social capital indicators are associated with political discussion on Yik Yak. Building upon previous research, we were specifically interested in exploring three contextually situated factors linked to campus-wide social capital formation.

The first factor of interest is *racial/ethnic heterogeneity*. Prior research has generally conceptualized high levels of racial/ethnic heterogeneity as exerting an inhibitory effect on

collective organization and action (e.g., Alesina & Ferrara, 2002; Vargo & Hopp, 2015). The theory driving this perspective is that "similarity breeds connections" (McPherson, Smith-Lovin, & Cook, 2001, p. 415) and; as such, *high* levels of racial/ethnic heterogeneity are thought to be indicative of comparatively *low* levels of social capital potential. However, on university campuses racial heterogeneity has been associated with a number of positive outcomes related to on-campus democratic engagement (e.g., Guren, Nagda, & Lopez, 2004).

The second factor of interest is *organizational participation*. Previous research has conceptualized social capital in terms of organizational participation (e.g., Putnam, 2000; Wellman, Hasse, Witte, & Hampton, 2001). Theoretically, associational involvement in the form of participation in religious, civic, voluntary, and interest oriented groups help encourage the development of shared norms and mutual trust (Portes, 1998).

Finally, we explored the degree to which participation in Greek life was a part of each university's social fabric. Participation in Greek organizations tends to "discourage interaction across difference" (Laird, 2005, p. 373). Specifically, involvement in Greek life has been conceptualized as an indicator of college students' involvement in "activities that do not challenge students to engage diverse peoples and ideas" (p. 373). On one hand, campuses featuring high levels of Greek involvement may feature heightened levels of social capital. Greek life forms strong, bonded networks between similar individuals. On the other hand, it may be the case that Greek involvement results in an insular social environment where the political and social striations necessary for democratic communication are suppressed.

Given the above literature, our third research question was interested in exploring if (and to what degree) contextual, campus-level characteristics (university demographics, university

learning environment indicators, and university social environment factors) are associated with political communication on Yik Yak:

RQ3: Do university demographics, university learning environment factors, and university social environment factors correlate with the frequency of political discussion on Yik Yak?

Method

Retrieving the Data from Yik Yak

Yik Yak is a freely available social media platform like Twitter. However unlike Twitter, it does not provide documentation for its Application Programming Interface (API). As such it is difficult to make standardized requests for data. This restriction is likely because of its policy to not to allow third-party apps and services to connect to it (Yik Yak, 2015). APIs are primarily used for these reasons. Computer science scholars have recognized this limitation and developed a solution to fetch Yik Yak data using python packages. The tool "YikYakTerminal" was used. It is an open source and freely available python program on GitHub (YikYakTerminal, 2015). It enables users to enter a custom location and retrieve messages. The authors here altered the code to automate data collection. The program ran every 24 hours and stored data locally in a tabular format. Yaks that were at least 4 hours old were collected each day. To classify each message's geographic location, each version of the script specified the "geofence" that corresponded to the university of interest. In all, 64 versions of the script were generated, one for each full time member in the Atlantic Coastal, Big Ten, Big 12, Pacific-12 and Southeastern conferences. The decision to focus on the above-described schools was made because (1) all of the selected universities are similar in terms of institutional structure; (2) all schools compete in high profile college athletics, which is a significant part of the student experience and thus likely to comprise

a substantial part of on-campus discussion; and (3) examination of approximately 60 contextual units is consistent with the number of clusters used by other researchers in similar big data analyses (e.g., Murthy, Gross, Pensavalle, 2015).

In the current study, data from 63 universities was analyzed. Data from the University of North Carolina was dropped because the geocoder incorrectly resolved the University of North Carolina at Chapel Hill to the Greensboro campus. This was because the term "the University of North Carolina" resolved to the latter on Google Maps. In all, 1,139,724 messages were acquired. Given that Yik Yak was in different stages of diffusion at each university, the number of messages per school ranged from 492 (Texas A&M) to 34,888 (University of Michigan). The average number of collected Yaks per university was 18,090.86 (SD = 9,058.16).

Coding for Political Talk

The researchers tried several approaches to extract political talk. First, the authors performed many iterations of Latent Dirichlet allocation (LDA) modeling to extract the topics inside of the Yik Yak data. The authors finally settled on Dynamic Modeling via Non-negative Matrix Factorization (Green & Cross, 2015). The authors adopted this LDA method because of its scalability to large datasets and its temporal nature. Standard topic modeling approaches assume the order of documents does not matter, making them unsuitable for time-stamped corpora. In contrast, *dynamic topic modeling* approaches track how language changes and topics evolve over time. Here we applied a two-level approach for dynamic topic modeling via Non-negative Matrix Factorization (NMF), which links together topics identified in snapshots of text sources appearing over time (Greene & Cross, 2016).

The authors extracted the top 100 topics, the top 15 of which can be seen in **Table 1**. As seen, while these topical models provided some insight into the types of content posted on Yik Yak, no clear political topics emerged.

INSERT TABLE 1 ABOUT HERE

To better hone in on political talk, the researchers adopted a lexicon approach. The researchers leveraged existing wordlists from a previous study, which detected political talk on Twitter from the 2012 election (Vargo & Guo, 2015). This list included 16 key political issues.[1] The lexicons originated from words that were traditionally associated with key political issues. In addition to the use of pre-existing wordlists, additional wordlists representing (1) words/phrases used to signify discussion of the State of the Union address (e.g., "State of the Union," "SOTU"); (2) key themes from the State of the Union Address (e.g., "NATO", "McConnell"); and (3) words/phrases related to current events (e.g., "ISIS," "Charlie Hebdo") were created.[2]

Like inter-coder agreement with traditional content analysis, "algorithms and dictionaries must often be repeatedly revised and tweaked to improve their performance" (Zamith & Lewis, 2015, pg. 4). This process is only complete when a satisfactory level of construct validity is established. In this case, two coders should agree with each other to establish "gold standard" data. Then, the computer should agree with the gold standard data at an acceptable level. The researchers conducted three rounds of validity checks. Wordlists were altered at each iteration to improve performance. Words that caused false positives were removed and more specific uses of the word were incorporated, or the word was removed. In cases where neither solution yielded more valid results, an exclusion list for that word was created. Words were also added to dictionary lists when researchers noticed their absence. At each stage, two human coders read a random sample of messages and coded them as either political or nonpolitical in nature. Then,

they compared their result with the computer-coded result. If the results did not match, the human coders reported an error. At all three stages the two human coders' intercoder agreement was acceptable ($\alpha > .85$). The two human coders agreed on whether talk was political or not with an acceptable degree of intercoder reliability.

For comparing the computer-coded data to the human-annotated data, percent agreement was used as an external validity check. A reliability check was not needed, because computers are presumed to be reliable (Riffe, Fico & Lacy, 2014). For the first iteration, the researchers examined a large subsample to explore the dataset and rigorously augment words. Subsequent sample frames were smaller, but matched the sample size standards set by previous research (Riffe, Fico & Lacy, 2014). The pairwise agreement for each round was as follows: Round 1 = 35.5% (n = 1,000); Round 2 = 73.4% (n = 282); Round 3 = 94.3% (n = 282).

University Data

U.S. News & World Report (USNWR) annually ranks U.S. post-secondary institutions. As a part of those rankings, USNWR analysts collect data points on 1,400 colleges and universities. The rankings have been compiled and published annually since 1985 and are the most widely quoted of their kind in the United States (Leiby, 2014). Data is collected via an annual survey sent to each school. There are several individual measures derived from this data.

University Demographics. To measure *university size*, USNWR data describing the estimated number of undergraduates enrolled at each university was used (M = 24,042.62, SD = 9,650.95). Here, we used undergraduate enrollment rather than total enrollment because Yik Yak is primarily used by undergraduate students (e.g., Huet, 2014; Northcut, 2015). For *attendance cost*, cost of attendance was used as the closest proxy for each university's relative socio-

economic status. The cost of attendance was calculated as the total fulltime tuition cost for a single year of attendance. For public schools, we used the in-state rate (M = $16, 863.27, SD = $13, 785.78). In terms of *public/private status*, 17.46% (*n* = 11) of the universities were private institutions. To measure the *gender distribution* at each university, we used the percentage of undergraduate male students (M = 50.21%, SD = 4.46). The average *age* of the undergraduate population at each school was 20.84 years (SD = 0.93 years). The *racial/ethnic distribution* at each university was measured as the percentage share of each university that was Asian/Asian-American (M = 8.25%, SD = 7.66%), Black/African-American (M = 5.87%, SD = 3.39%), Hispanic/Latino (M = 8.77%, SD = 5.96%), and White/Caucasian (M = 64.80%, SD = 13.64%).

University Learning Environment. Two measures were used evaluate *average classroom size*. The first measure was the percentage of classes at each university with less than 20 students (M = 42.45%, SD = 11.66%). The second measure was the percentage of classes at each university with greater than 50 students (M= 15.95%, SD = 5.26%). *Graduation rate* was measured as the percentage of students who graduate in six years or less (M = 75.73%, SD = 11.43%). *Freshman retention rate* was calculated as the percentage of first year students who return for a second year (M= 89.32%, SD = 5.77%).

University social environment. As it pertains to *campus heterogeneity,* the USNWR data provides the following racial/ethnic/identification categories: Black/African-American, American Indian/Native American, Asian/Asian-American, Hispanic/Latino, White/Caucasian, Pacific Islander, 2 or more races, International, and unknown. To generate an estimate of racial/ethnic heterogeneity on each campus, we employed the approach previously used by Costa & Kahn (2003). Use of this technique returned a value between 0 and 1, wherein higher values were indicative of greater levels of on-campus heterogeneity (M= 0.53; SD = 0.13).[3] Next, we

measured the number of *university-affiliated organizations* using the USNWR's pooled estimate of number of campus wide organizations (M = 613.83, SD = 290.24). Finally, we pooled the number of fraternities and sororities to create a single index measuring *campus-wide Greek involvement* (M = 38.34, SD = 21.21).

Results

The first research question addressed the degree to which Yik Yak featured political talk. Over the entire month of January, we observed 9,340 political comments. As such, political comments comprised less than 1% of all comments posted on Yik Yak (9,340 political messages/1,139,724 total messages = 0.82%). Averaged across all universities, the mean percentage of political talk was 0.84% (SD = 0.45%). Examination of the individual university totals suggested that Duke University's Yik Yak featured the lowest overall percentage of political discussion (0.29%) while the University of Kansas' Yik Yak featured the highest overall percentage of political discussion (3.41%).[4] percentage of political discussion at each university is shown in Figure 1.

INSERT FIGURE 1 ABOUT HERE

The second research question was concerned with the degree to which discussion on Yik Yak is responsive to or reflective of national political events. January 20th, the day of the State of the Union Address, saw the highest overall percentage of political discussion on Yik Yak (2.63%; M = 2.59%, SD = 2.24%). As seen in **Figure 2**, Yik Yak-based political talk was substantially elevated during the four-day period beginning on January 19. 16.10% percent (*n* = 1,504) of all observed political comments were posted on January 20th and 36.26% of all observed political comments (*n* = 3,387) were posted during the four-day period from January 19 to January 22.

On January 20th, the day of the State of the Union address, individual percentages for each university indicated that political talk had substantive variance, ranging from 0.00% (Texas A&M, University of Arizona, University of Texas) to 14.8% (University of Kansas). Five university accounts featured single day averages in excess of 5.0% (University of Kansas = 14.80%, Northwestern University = 9.51%, University of Virginia = 7.18%, Boston College = 5.84%, Purdue University = 5.48%, Baylor University = 5.11%). The university with the greatest single-day percentage of political talk was the University of Kansas (21.05%; January 22).

INSERT FIGURE 2 ABOUT HERE

Next, we explored the degree to which political discussion on January 20th differed from "normal" levels of onsite political discussion. Paired samples *t*-tests suggested that the mean percentage of political talk on January 20th was significantly higher than the percentage of political talk observed on January 19th, $t(62) = 5.94$, $p < .01$, January 21st, $t(62) = 5.59$, $p < .01$, and January 22nd, $t(62) = 5.88$, $p < .01$. Given that January 19, 21, and 22 featured the three next highest percentages of political conversation (relative to January 20), these findings allow us to conclude that the percentage of political conversation on January 20th was statistically greater than on any other day in the period of observation.

Finally, the degree to which higher levels of political talk during "normal" periods was associated with political talk during the four day period surrounding the State of the Union address (January 19 – January 22). To do so, two data points were calculated for each university. The first series of data points represented the percentage of all talk that was political for the periods between January 1 – January 18 and January 23 – January 31. The second series of data points represented the percentage of all talk that was political for the period between January 19 and January 22. Using these values, we estimated a bivariate correlation describing the

relationship between "normal" levels of political communication and levels of political communication surrounding the State of the Union address. The results suggested that universities that featured higher levels of political communication during "normal" periods of the month also featured heightened levels of political communication during the four-day period bracketing the State of the Union address, $r = .28$, $p < .05$.

The third research question was interested in exploring the degree to which university characteristics were associated with the prevalence of political discussion on Yik Yak. To explore this question, we examined the bivariate correlations between the percentage of political discussion on each campus and university-level indicators describing university demographics, learning environment, and social capital potential. Given that a bulk of the observed political communication occurred between January 19 and January 22 and that the nature of that talk was reflexive of a special event, correlations were estimated for both the full month percentage and the percentage for the period surrounding the State of the Union address. For the university demographic factors, we observed significant, negative correlations between university size and the amount of political talk for both the full month ($r = -.30$, $p < .05$) and the 4-day period surrounding the State of the Union address ($r = -.28$, $p < .05$). For the learning environment factors, there was a ~~positive~~negative and ~~,~~ significant correlation between the percentage of classes with > 50 students and the amount of political discussion for both the full month ($r = -.31$, $p < .01$) and the period between January 19 and January 22 ($r = -.32$, $p < .01$). Finally, as it relates to university social environment, our results suggest that campuses that feature a high number of Greek organizations were less likely to feature political discussion. Specifically, we observed negative correlations between number of on-campus Greek organizations and the frequency of political discussion for both the month-long and four-day period from January 19 to

January 22 (r = -.32, p < .05 for both periods). See Table 2 for a full summary of the observed correlations.[5]

INSERT TABLE 2 ABOUT HERE

Discussion

Political Talk Overall

Political discussion is increasingly located online. The rapid diffusion and adoption of social media has the ability to re-shape how, where, and when people talk about important political and social issues. However, as scholars have previously pointed out, the democratic potential of ~~social media~~online platforms are not universally realized (e.g., Papacharissi, 2004). The current study explored the degree to which young adults use Yik Yak for political expression. To the best of our knowledge, this study is a number of firsts as it pertains to Yik Yak. It is the first to address political communication, the first to leverage computational social science methods with big data, and the communication sciences to explore the uses of Yik Yak.

Political engagement on Yik Yak appears to be low. Less than 1% of all observed comments were political in nature. However, evidence suggests that the volume of Yik Yak-based political -commentary is responsive to highly visible major political events. The data indicated that 36.26% of all observed political comments were posted in the four days surrounding President Obama's 2015 State of the Union address. This is similar to what previous researchers have found for Twitter (e.g., Vargo, Shaw, & Basilaia, 2015). -

Finally, the data suggested that political conversation on Yik Yak was lower at universities with large undergraduate enrollments, a large percentage of classes with greater than 50 students, and a large number of fraternities. Implications are discussed below.

LDA topic modeling (see Table 1) indicated that Yik Yak is dedicated to interpersonal relationships, sexual topics, and school-related topics. Political talk is completely absent from the most popular types of conversations. Yik Yak's potential for political discussion is primarily unrealized. However, Yik Yak is responsive to major political events. A large proportion of the observed comments were posted on or around the day of the State of the Union address. The results suggested a positive, linear relationship between amount of "normal" monthly political talk and amount of talk that occurred during the four-day period surrounding the address. This suggests some schools are more attentive to politics than others, and this effect is boosted by major events. These findings provide scholars and practitioners a baseline upon which to build more sophisticated predictive models of college students' political engagement on Yik Yak. Schools' relative political "baseline" seems to be a clue as to how it will respond during events.

The Technological Affordances of YikYak

Anonymity. Yik Yak is a unique in that it provides ability for users to be anonymous to other users. This is a technological affordance undoubtedly changes the nature of the discussion on the medium (see Table 1). This change could allow for lesser barriers to political discussion. For instance, individuals could voice divergent opinions without fear of personal consequence. Unlike Facebook and Twitter, political messages that are controversial would not spill over into the identities of real people. Moreover, it is well-known that trolls, or individuals who deliberately post offensive or provocative online posting with the aim of upsetting someone, prefer anonymity (Phillips, 2015). Trolls are known to post divergent political opinions (Hmielowski, Hutchens and Cicchirillo, 2014). For these reasons, it is interesting that YikYak has not afforded more political talk, whether it civil or ~~u~~incivil in nature.

Instead, Table 1 shows that anonymity in a college setting appears to encourage more discussion about sex, drinking alcohol and consuming illicit substances (e.g., marijuana and other drugs). Anonymity still plays a key role in these messages, as these behaviors can be illegal for college students. The reasons why these behaviors are discussed and political discussion is relatively absent is ripe for further study. At present, no known studies have tackled the uses and gratifications of YikYak.

Geolocations. Beyond privacy, YikYak also affords geofenced content. In one sense this is a novel feature of the platform. Users of YikYak can only see and post messages that correspond to their geographical area. Here, this design choice makes it easy for users to see the happenings of only their local area. This locus also seems to encourage discussion of events that pertain to those specific areas (e.g., here a specific college campus). Table 1 again reveals that the daily happenings of college students is prevalent in the data. Keywords that relate to campus life (e.g., classes and professors) appear often. Location-based messaging may itself also discourage certain types of messages. For instance, Twitter is a medium that does not filter data via geographical areas. Instead, all users can see the entire population of Twitter content. Twitter is also immensely popular during national events (e.g., a presidential debate) and usage of the platform spikes during these times (Hu et al., 2012). Here, we studied the State of the Union, a national event that does typically generate significant traffic on Twitter (Wang et al., 2012). It could be that the intentional design choice for YikYak to filter messages locally is itself quelling discussions that are more global in nature, instead giving preference to events that are more local in nature. Further research should examine politics at local levels to see if YikYak and other geofenced social media are more responsive in these situations.

Contextual Factors

~~Second, t~~The data tentatively suggested that contextual factors may have an influence on the frequency to which political discussion appears on Yik Yak. This is consistent with previous research that suggests similar interplays between online and offline behaviors (e.g., Dahlgren, 2000). Campuses with a high number of classes that had more than 50 students were less likely to feature high levels of political communication on Yik Yak. This finding could be due to a conditioning effect. Students in larger classrooms are given fewer opportunities for classroom discussion. Thus, they may be less likely to use their free time to discuss such issues. Similarly, our observation that the number of Greek organizations was negatively affiliated with political discussion could indicate that Greek associations institute social structures that are non-conducive to political talk. [6]

It is also important to note that we observed a large number of null relationships between the contextual factors of interest and the frequency of political posts on Yik Yak. One explanation for these findings could be the relative homogeneity of the modern university environment. Another explanation could be moderating factors that exert conditional influences on Yik Yak-based political behavior. As an emergent platform, Yik Yak could also be unevenly diffused across universities. This could mask potentially meaningful relationships. Clearly, in this area, future research is needed.

This study has a number of limitations. We did not address the quality of political communication on Yik Yak. A comprehensive understanding of any media's democratic potential must encompass both *quantity* and *quality* (Hopp & Vargo, 2015). As such, future research could investigate the degree to which political discussion on Yik Yak is free of name-calling, insults, and other indicators of low quality political discussion. Additionally, given the application's localized nature, it may be the case that Yik Yak is more frequently used to discuss

local, rather than national, political, social, and cultural issues. Future research could comparatively explore the degree to which Yik Yak is used for discussion of local political issues. Third, the structure of the USNWR data framed the operationalization of some of the campus-level variables used in this study. In some cases, this may have resulted in measures with less-than-ideal properties or restricted us from exploring potentially meaningful relationships. Future research should seek to employ more robust contextual measures as a means of substantiating and building upon the current results.

Twitter and Facebook have played a prominent role in the discussion of the last two elections, especially among young adults. Campaigns have reacted to this new reality by building sophisticated social media campaign strategies aimed at building and sustaining candidate-friendly narratives. Given the rapid adoption of Yik Yak among young voters, it seems entirely conceivable that platforms such as Yik Yak may play a role in future elections. As such, we hope that scholars build upon this initial work in order to build a fuller understanding of Yik Yak's democratic potential.

Footnotes

1. The 16 ongoing issues frequently discussed in political context are: (1) taxes; (2) jobs/unemployment; (3) federal budget deficit; (4) general economy; (5) foreign affairs; (6) immigration; (7) healthcare; (8) public order; (9) LGBT/same-sex marriage; (10) abortion; (11) environment/climate; (12) energy; (13) education; (14) role of government; (15) middle class issues; and (16) welfare/entitlements

2. Complete lexicons and Python code used to classify the data are freely available from the authors.

3. The formula used to generate the heterogeneity measure was

$$1 - \sum_{k} s_{ki}^{2}$$

where s_{ki} is the share of race/ethnic/identification category k for university i.

4. Additional analyses suggested that the number of total university-wide messages was not related to the percentage of political messages, $r = -.02$, $p > .05$.

5. The USNWR data included some missing values. Specifically, 1 observation was missing for the gender breakdown measure, 2 observations were missing for the measure describing university-affiliated clubs/organizations, and 7 observations were missing for the measure describing number of on-campus Greek organizations. To determine the degree to which these missing values biased our results, we used the *R* package "mice" to impute missing values. Using these values, we estimated all relevant correlations. These correlations were then compared to the correlations acquired using listwise deletion. In all cases, the relational magnitude and significance levels were essentially identical, suggesting that missing values did not bias our reported results.

6. Given the exploratory nature of the current study, the relatively small number of campus units, the presence of missing data, and the strong correlations between many of the contextual indicators, we chose not to use regression analyses to model the current data. However, when the three significant factors identified in Table 2 (undergraduate enrollment, class sizes > 50, and number of Greek organizations) were simultaneously entered into a regression equation, class size emerged as the strongest overall predictor, followed by number of Greek organizations and undergraduate enrollment.

References

Alesina, A., & Ferrara, E. L. (2000). Participation in heterogeneous communities. *Quarterly Journal of Economics, 115,* 847 – 904.

Blanchard, A., & Horan, T. (1998). Virtual communities and social capital. *Social Science Computer Review*, 16, 293 – 307.

Boardman, J. D., & Robert, S. A. (2000). Neighborhood socioeconomic status and perceptions of self-efficacy. *Sociological Perspectives*, *43*(1), 117-136.

Braxton, J. M., & McClendon, S. A. (2001). The fostering of social integration and retention through institutional practice. *Journal of College Student Retention, 3,* 57 – 71.

Chaffee, S. H., Ward, S. L., & Tipton, L. (1981). Mass communication and political socialization. In M. Janowitz & P. Hirsch (Eds.), *Reader in public opinion and mass communication* (pp.74–88). New York: Free Press.

Chamlin, M. B., & Cochran, J. K. (1995). Assessing Messner and Rosenfeld's institutional anomie theory: A partial test. *Criminology, 33,* 411 – 429.

Conner, C. (2015). Activist groups ask the Department of Education to address Yik Yak harassment. *Student Press Law Center.* From http://www.splc.org/blog/splc/2015/10/activist-groups-ask-the-department-of-education-to-address-yik-yak-harassment

Correa, T., & Jeong, S. H. (2011). Race and online content creation: Why minorities are actively participating in the web. *Information, Communication, & Society, 14,* 638 – 659.

Costa, D. L., & Kahn, M. E. (2003). Civic engagement and community heterogeneity: An economist's perspective. *Perspectives on Politics, 1,* 103–111.

Dahlgren, P. (2000). The Internet and the democratization of civic culture. *Political Communication, 17*, 335 – 340.

Diddi, A., & LaRose, R. (2006). Getting hooked on the news: Uses and gratifications and the formation of news habits among college students in an Internet environment. *Journal of Broadcasting & Electronic Media, 50,* 193 – 210.

Duggan, M., Ellsion, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). Social media update 2014. *Pew Research Center*. From: http://www.pewinternet.org/files/2015/01/PI_SocialMediaUpdate20144.pdf

Ellison, N. B., Steinfeld, C., & Lampe, C. (2007). The benefits of Facebook "Friends:" Social capital and college students use of online asocial network sites. *Journal of Computer-Mediated Communication, 12,* 1143 – 1168.

Gil de Zúñiga, H., Molyneux, L., & Zheng, P. (2014). Social media, political expression, and political participation: Panel analysis of lagged and concurrent relationships. *Journal of Communication, 64,* 612 – 634.

Glascock, J. (2014). Contribution of demographics, sociological factors, and media usage to verbal aggressiveness. *Journal of Media Psychology*.

Greene, D., & Cross, J. P. (2016). Exploring the Political Agenda of the European Parliament Using a Dynamic Topic Modeling Approach. *arXiv preprint arXiv:1607.03055*.

Guo, L., & Vargo, C. (2015). The Power of Message Networks: A Big-Data Analysis of the Network Agenda Setting Model and Issue Ownership. *Mass Communication and Society*, *18*, 557-576.

Hargittai, E., & Litt, E. (2011). The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults. *New Media & Society, 13,* 824Media

Hendryx, M. S., Ahern, M. M., Lovrich, N. P., & McCurdy, A. H. (2002). Access to health care and community social capital. *Health Services Research, 37,* 85 – 101.

Hess, A. (2015). Don't ban Yik Yak. *Slate*. From http://www.slate.com/articles/technology/users/2015/10/yik_yak_is_good_for_university_students.html

Hillygus, D. S. (2005). The missing link: Exploring the relationship between higher education and political engagement. *Political Behavior, 27*, 25 – 47.

Hmielowski, J. D., Hutchens, M. J., & Cicchirillo, V. J. (2014). Living in an age of online incivility: Examining the conditional indirect effects of online discussion on political flaming. *Information, Communication & Society, 17*, 1196–1211.

Hu, Y., John, A., Seligmann, D. D., & Wang, F. (2012, May). What Were the Tweets About? Topical Associations between Public Events and Twitter Feeds. In *ICWSM*.

Huet, E. (2014). Yik Yak wants to be a news hub but it needs to grow up first. *Forbes.* From http://www.forbes.com/sites/ellenhuet/2014/10/20/yik-yak-wants-to-be-a-news-hub-but-it-needs-to-grow-up-first/

Huysman, M., & Wulf, V. (2004). *Social capital and information technology.* Cambridge, MA: MIT Press.

Karp, D. A., & Yoels, W. C. (1976). The college classroom: Some observations on the meanings of student participation. *Sociology & Social Research, 60*, 421 – 439.

Kwak, N., Williams, A. E., Wang, X., & Lee, H. (2005). Talking politics and engaging politics: An examination of the interactive relationship between structural features of political talk and discussion engagement. *Communication Research, 32,* 87 – 111.

Kuh, G. D. (1993). In their own words: What students learn outside the classroom. *American Educational Research Journal*, 30, 277 – 304.

Laird, T. N. (2005). College students' experiences with diversity and their effects on academic self-confidence, social agency, and disposition toward critical thinking. *Research in Higher Education, 46,* 365 – 387.

Lau, L. K. (2003). Institutional factors affecting student retention. *Education, 124,* 126 – 136.

Lee, J. & Lee, H. (2010). The computer-mediated communication network: Exploring the linkage between online community and social capital. *New Media & Society, 12,* 71 – 727.

Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., & Smith, A. (2013). Teens, social media,
and privacy. Pew Internet a nd A merican Li fe Project.
Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., & Smith, A. (2013). Teens, social media,
and privacy. Pew Internet a nd A merican Li fe Project.
Leiby, R. (2014, September 9). The guru behind the U.S. News college rankings. *The Washington Post.* Retrieved November 9, 2015, from https://www.washingtonpost.com/lifestyle/style/the-us-news-college-rankings-guru/2014/09/09/318e3370-3856-11e4-8601-97ba88884ffd_story.html

Lizzio, A., Wilson, K., Simons, R. (2002). University students' perceptions of learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education, 27,* 27 – 52.

Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., & Smith, A. (2013). Teens, social media, and privacy. *Pew Internet and American Life Project.*

McClintock, C. G., & Turner, H. A. (1962). The impact of college upon political knowledge, participation, and values. *Human Relations, 15,* 163 – 176.

McPherson, M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review, 52,* 370-389.

Murthy, D., Gross, A., & Pensavalle, A. (2015). Urban social media demographics: An exploration of Twitter use in major American cities. *Journal of Computer Mediated Communication.* Online before print. doi: 10.1111/jcc4.12144

Newton, K. (2003). Trust, social capital, civil society, and democracy. *International Political Science Review*, 22, 201 – 214.

Onyx, J., & Bullen, P. (2000). Measuring social capital in five communities. *The Journal of Applied Behavioral Science, 36,* 23 – 42.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society, 6,* 259 – 283.

Phillips, Whitney (2015). *This is why we can't have nice things. Mapping the relationship between online trolling and mainstream culture*. Cambridge, MA: The MIT Press.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology, 24,* 1 -24.

Pretty, J., & Ward, H. (2001). Social capital and the environment. *World development*, *29*(2), 209-227.

Pulvers, K., & Diekhoff, G. M. (1999). The relationship between academic dishonesty and college classroom environment. *Research in Higher Education, 40*, 487 – 498.

Putnam, R. (2000). *Bowling alone; The collapse and revival of the American community*. New York: Simon & Schuster.

Riffe, D., Lacy, S. & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research*. New York: Routledge/Taylor & Francis Group.

Rosenfeld, R., Messner, S. F., & Baumer, E. P. (2001). Social capital and homicide. *Social Forces, 80,* 283 – 309.

Rupasingha, A., Goetz, S. J., & freshwater, D. (2006). The production of social capital in US counties. *The Journal of Socio-Economics, 35,* 83 – 101.

Shah, D. V. (1998). Civic engagement, interpersonal trust, and television use: An individual level assessment of social capital. *Political Psychology, 19,* 469 – 496.

Shah, D. V., Cho, J., Eveland, W. P., & Kwak, N. (2005). Information and expression in a digital age: Modeling Internet effects on civic participation. *Communication Research, 32*, 531 – 565.

Shah, D. V., Kwak, N., & Holbert, L. (2001). "Connecting" and "disconnecting" with civic life: Patterns of Internet use and the production of social capital. *Political Communication, 18,* 141 – 162.

Wellman, B., Haase, A. Q., Witte, J., & Hampton, K. (2001). Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American Behavioral Scientist, 45*, 436 – 455.

Valenzuela, S., Kim, Y., & Gil de Zúñiga, H. (2012). Social networks that matter: Exploring the role of political discussion for online political participation. *International Journal of Public Opinion Research, 24,* 163 – 184.

Valenzuela, S., Park, N., & Kee, K. F. (2009). Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation. *Computer-Mediated Communication, 14,* 875 – 901.

Vargo, C. J., & Hopp, T. (2015). Socioeconomic status, social capital, and partisan polarity as predictors of political incivility on Twitter: A congressional district-level analysis. *Social Science Computer Review.* doi: 10.1177/0894439315602858

Vargo, C., Basilaia, E. & Shaw, D. (2015). Event vs. Issue: Twitter Reflections of Major News, a Case Study. ***Studies in Media and Communication, 9, 215-239.***

Wang, H., Can, D., Kazemzadeh**, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In Proceedings of the ACL 2012 System Demonstrations (pp. 115-120). Association for Computational Linguistics.**

Yamamoto, M., Kushin, M. J., & Dalisay, F. (2015). Social media and mobiles as political mobilization forces for young adults: Examining the moderating role of online political expression in political participation. *New Media & Society, 17,* 880 -898.

YikYakTerminal on GitHub. (2015, November 4). Retrieved from https://github.com/djtech42/YikYakTerminal

Yik Yak Terms of Service. (2015, November 4). Retrieved from http://www.yikyakapp.com/terms/

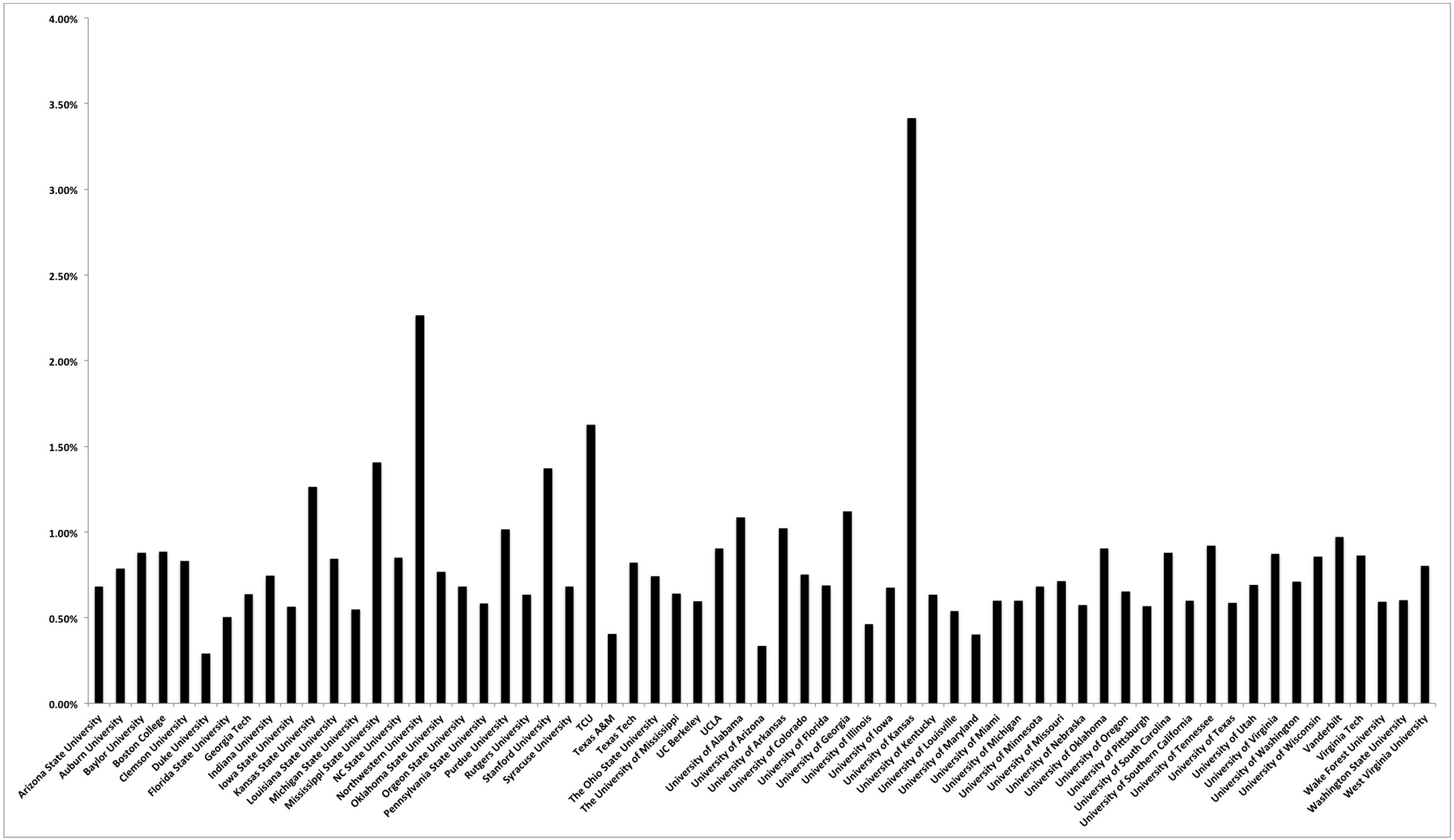

Figure 1. Percentage of political talk on Yik Yak for each university in the sample

Note: Reported percentages are for the entire month of January 2015 ~~(M = 0.82, SD = 0.45%~~

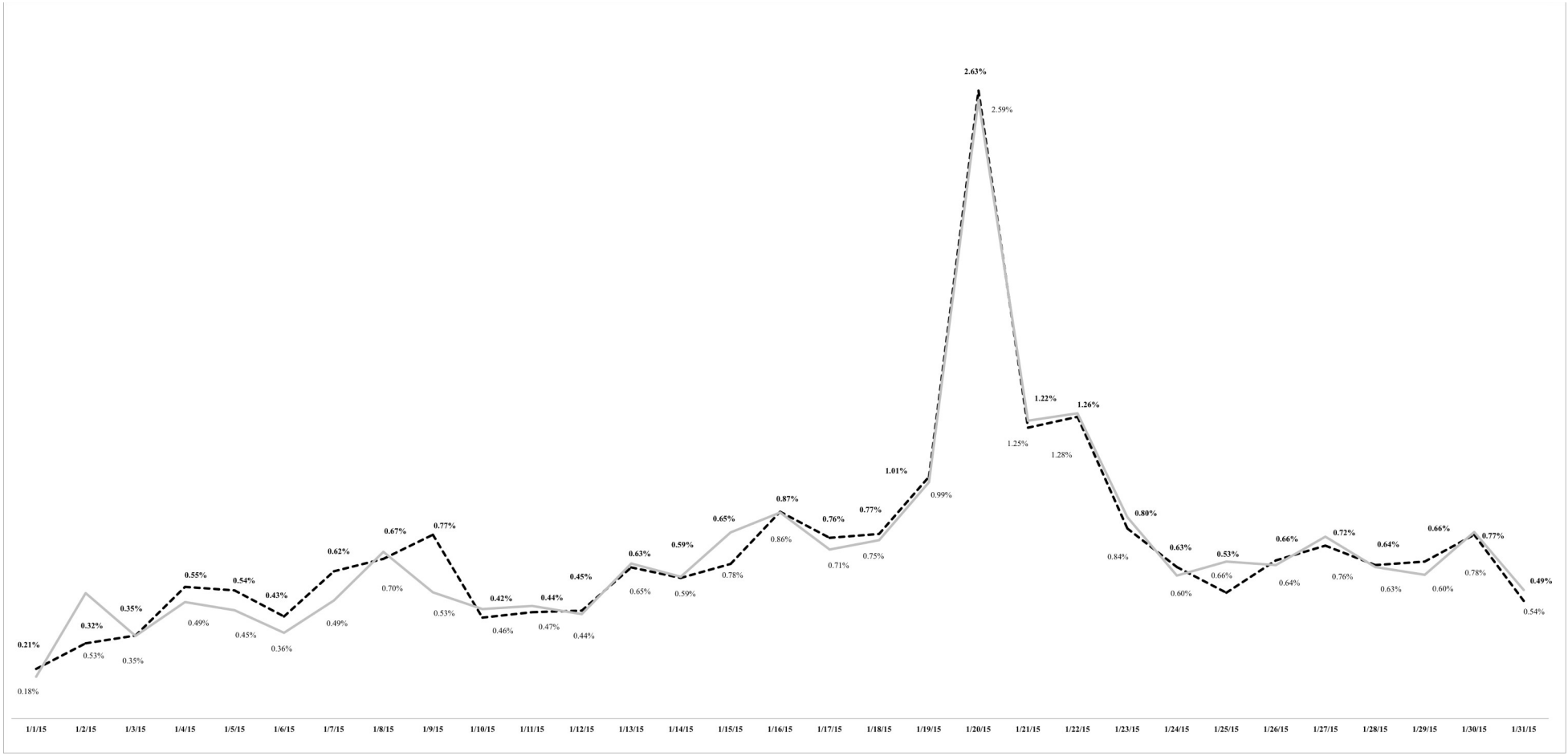

Figure 2. *Per-day summary of Yik Yak political talk.*

Note: The dashed black line represents the raw percentage of political talk per day (total number of political comments / total number of comments). The solid gray line represents the average per-day percentage of political talk for all sampled universities. Values for the raw percentage of daily political talk are placed above the dashed black line while the averaged values are placed below the dashed gray line.

Table 1

Top 15 LDA topics

Topic Number	Topic Contents
1	*like look looks feels smells sounds act feeling doesn smell women days bitch person actually*
2	*just saw wanna realized said wish walked mean took did told tell doesn woke watched*
3	*don understand care anymore wanna tell worry talk wear need text mind work trust drink*
4	*people white actually black stop talk understand meet world lot use talking wonder tinder yaks*
5	*girl cute hot scout sorority ask date beautiful looking white saw cookies talk wants said*
6	*day valentine valentines snow single leg great having today beautiful classes mlk gonna long ll*
7	*fuck shut buddy yeah state did bu bitch uva wanna holy drunk gonna bitches floor*
8	*class professor minutes tomorrow late 30 walk early semester teacher morning cancelled sitting*
9	*time long spend start nap remember half waste having work hard actually phone free spent*
10	*guy cute hot saw nice black talking date looking friend tell wants said wearing ask*
11	*know didn doesn ll anybody real did person dont bad doing gonna ya won talking*
12	*girls sorority white hot talk tinder cute wear looking boys attractive black beautiful pretty*
13	*today need help stop didn open saw classes wearing work gonna weather buddy man did*
14	*feel bad makes better person way lonely making sad guilty sick doing weird sorry boyfriend*
15	*going tomorrow home gym semester weekend start instead im hell break alarm today outside*

Table 2

Bivariate correlations between percentage of political talk on Yik Yak and university characteristics

	Political Talk January 1 -31	**Political Talk January 19-22**
Demographics		
University Size	-.30**	-.28**
Attendance Cost	.13	.15
Private Status	.16	.16
Gender (Percent Male)	-.10	-.12
Average Age	-.07	-.06
Percentage Asian	-.09	-.09
Percentage Black	.05	.00
Percentage Hispanic	-.11	-.08
Percentage White	.12	.10
Learning Environment		
Percent of Classes < 20 Students	.22*	.17
Percent of Classes > 50 Students	-.31***	-.32***
Graduation Rate	-.09	-.10
Freshman Retention	-.13	-.14
Social Environment		
Campus Heterogeneity	-.13	-.09
Number of University-Affiliated Clubs	-.17	-.17
Number of Greek Organizations	-.32**	-.32**

Note: * $p < .10$, ** $p < .05$, *** $p < .01$