Communication Faculty Publications

Communication

1-1-2018

Beyond Big Bird, Binders, and Bayonets: Humor and Visibility Among Connected Viewers of the 2012 US Presidential Debates

Kevin Driscoll
University of Virginia, kdriscoll@virginia.edu

Alex Leavitt
Facebook Research

Kristen L. Guth
University of Kentucky, Kristen.Guth@uky.edu

François Bar
University of Southern California

Aalok Mehta
University of Southern California

Repository Citation

Beyond Big Bird, Binders, and Bayonets: Humor and Visibility Among Connected Viewers of the 2012 US Presidential Debates

Notes/Citation Information

Published in *Social Media + Society*, v. 4, issue 1, p. 1-12.

Beyond Big Bird, Binders, and Bayonets: Humor and Visibility Among Connected Viewers of the 2012 US Presidential Debates

Kevin Driscoll[1], Alex Leavitt[2], Kristen L. Guth[3], François Bar[4],
and Aalok Mehta[4]

Abstract

During the 2012 US presidential debates, more than five million connected viewers turned to social media to respond to the broadcast and talk politics with one another. Using a mixed-methods approach, this study examines the prevalence of humor and its relationship to visibility among connected viewers live-tweeting the debates. Based on a content analysis of tweets and accounts, we estimate that approximately one-fifth of the messages sent during the debates consisted of strictly humorous content. Using retweet frequency as a proxy for visibility, we found a positive relationship between the use of humor and the visibility of individual tweets. Not only was humor widespread in the discourse of connected viewers, but humorous messages enjoyed greater overall visibility. These findings suggest a strategic use of humor by political actors seeking greater shares of attention on social media.

Keywords

political communication, humor, social media, attention, live-tweeting

In 2012, political talk on Twitter peaked over three October evenings during which the leading candidates in the US presidential election, Democrat incumbent Barack Obama and Republican challenger Mitt Romney, appeared in televised debates (Hong & Nadler, 2012). Millions of connected viewers turned to Twitter to comment on the unfolding debates and read the comments of others (Bruns & Highfield, 2015; Christensen, 2013; Conway, Kenski, & Wang, 2013). This outpouring of political talk ranged from the phatic to the thoughtful, from the benign to the outrageous, but it was the humorous content—sarcastic, snarky, biting, and bizarre—that seemed to dominate the experience of "live-tweeting" the debates (Freelon & Karpf, 2015; Tsou et al., 2013).

Political humor figured prominently in the discourse surrounding the 2012 presidential debates, both in the real-time experiences of connected viewers and in the post hoc debate coverage published by news organizations. Critics at NPR, CNN, and others ranked the "funniest" tweets of each debate, promised to explain debate-related jokes such as "binders full of women," and entreated politicians to adapt to the new reality of social media parody (e.g., Dailey, 2012; Katz, 2012; Kelly, 2012). This anecdotal coverage of humor

was set against a backdrop of macroscale statistics and charts produced by third-party social media analytics services such as Crimson Hexagon and Topsy (Crimson Hexagon's Editorial Team, 2012; Sharp, 2012; vanessa, 2012). Neither the news organizations nor the commercial data analytics firms disclosed the overall volume of social media talk dedicated to humor, despite the well-known technical and analytic challenges posed by such polysemous political talk (Bakliwal et al., 2013; Gayo-Avello, 2012; Mejova, Srinivasan, & Boynton, 2013; Wang, Can, Kazemzadeh, Bar, & Narayanan, 2012). The present research was designed, therefore, to estimate the prevalence of humor among the tweets sent by connected viewers of the 2012 US presidential debates and evaluate the impact of

[1]University of Virginia, USA
[2]Facebook Research, USA
[3]University of Kentucky, USA
[4]University of Southern California, USA

Corresponding Author:
Kevin Driscoll, Department of Media Studies, University of Virginia, P.O. Box 400866, Charlottesville, VA 22903-1738, USA.
Email: kdriscoll@virginia.edu

humor on the visibility of particular voices within that live-tweeting population:

> *RQ1. How much of the overall connected viewing activity was dedicated to humor?*
> *RQ2. What was the relationship between humor and visibility among live-tweeting viewers?*

To assess the prevalence of humor among connected viewers, we undertook a mixed-methods analysis of a large corpus of Twitter messages sent on the nights of the three debates. Rather than attempt to identify specific types of humor, we adopted a broad definition of humor inclusive of a variety of genres and styles. After analyzing multiple large samples, we estimate that at least one-fifth of all tweets sent during the debates represented some form of humor. Furthermore, we found that humor contributed significantly to the visibility of individual tweets, as measured by their retweet counts. These findings provide a baseline for future research on social media, humor, and political communication. If humor affords greater visibility to political speech, then we have to get serious about jokes.

Literature Review

Live-Tweeting and Visibility

During a US presidential debate, elite and non-elite audiences alike may react and respond in real time using a social media system such as Twitter (Freelon & Karpf, 2015). This practice, known as "live-tweeting" (Schirra, Sun, & Bentley, 2014) or "connected viewing" (Holt & Sanson, 2013; Pittman & Tefertiller, 2015) is common during televised events such as award shows, popular dramatic series, and major sporting events, as well as political debates and speeches (Bentivegna & Marchetti, 2015; Highfield, Harrington, & Bruns, 2013; Houston, McKinney, Hawthorne, & Spialek, 2013; McKinney, Houston, & Hawthorne, 2014). Live-tweeting may take on more pronounced political significance during "breaking" events such as protests, rallies, and emergencies (Bruns, Burgess, Crawford, & Shaw, 2012; Papacharissi & Oliveira, 2012) when participation offers connected viewers the sense of "being a part of" an unfolding drama of historical significance (Thorson, Hawthorne, Swasy, & McKinney, 2015). Across these contexts, Twitter provides a common channel for the voices of everyday viewers, pundits, politicians, and news organizations.

Of course, very few tweets sent by everyday viewers were seen by millions; most remained obscure. On a social media system such as Twitter, where the barriers to expression are low, messages compete with one another for the scarce attention of potential audiences (Webster, 2014). In this context, "visibility" describes the size of a tweet's potential audience, a precondition for exposure and influence. The visibility of a tweet is shaped initially by contextual factors such as timing

and the author's prior notoriety (Hodas & Lerman, 2012; Pezzoni, An, Passarella, Crowcroft, & Conti, 2013) and later by the mechanics of the platform such as the sequencing of each user's timeline and the selection of Trending Topics (Bucher, 2012; Gillespie, 2012).

Twitter users actively participate in the allocation of visibility through the practice of "retweeting." Retweeting is a platform-specific form of recirculation, analogous to "sharing" on Facebook or "reblogging" on Tumblr. On an individual basis, retweeting is a conversational practice (Christensen, 2013) but taken in aggregate, retweeting offers a proxy measure of visibility (boyd, Golder, & Lotan, 2010). Zhang, Wells, Wang, and Rohe (2017) theorize retweeting as the "amplification" of visibility, a means of signaling that a particular tweet and its author are worthy of attention. These signals are continuously fed back into the Twitter platform as inputs to various ranking and recommendation algorithms that, in turn, contribute to the ongoing distribution of visibility.

Political Humor

Political humor is a particularly potent form of political communication, used to critique and expose politicians' values, ineptitude, abuses of power, or arrogance (Nilsen, 1990). In some instances, political humor has moved audiences to political action (Jones, Baym, & Day, 2012). In the early 2000s, political entertainment programs such as *The Daily Show* developed a new form of political humor combining elements of news, commentary, satire, and late-night comedy (Baym, 2005, 2007; Warner, 2007). Later iterations, such as *The Colbert Report*, presumed the connectedness of their audiences, extending engagement with viewers across social media and the Web (Burwell & Boler, 2014). While political humor has long been a component of mainstream political culture, researchers are only beginning to explore the circulation of political humor through social media.

Humor is fundamental to Internet-mediated communities and cultures (Shifman, 2007; Shifman & Blondheim, 2010). For many users, the decision to amplify a message on Twitter is based on a desire to share something meaningful, poignant, or entertaining (Jenkins, Ford, & Green, 2013). In their exploration of virality, Nahon and Hemsley (2013) found that humorous media "resonates" with viewers in a way that may motivate sharing. On Twitter specifically, humor connects journalists and politically interested participants (Holton & Lewis, 2011), and during live-tweeting events, tweets containing humor may reach greater visibility and longevity than breaking news messages (Highfield, 2015b). In addition, the structure of Twitter affords platform-specific forms of political humor, such as parodies of public figures (Highfield, 2015a).

Political humor, however common, remains elusive to social media researchers. Humor is difficult for both human and non-humor coders to reliably identify because of its linguistic variety and dependence on context (Attardo, 2014). Few researchers have tackled the problem of reliably
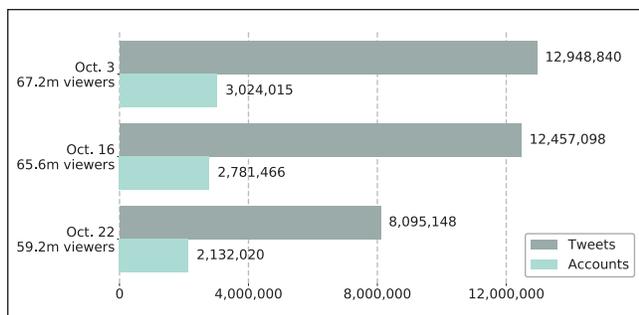
Figure 1. Connected viewership during the 2012 US presidential debates.
Note. Television audience estimates based on Nielsen (2012a; 2012b; 2012c).

identifying humor specifically in the context of Twitter. In a comparative study of sentiment analysis software for Twitter, joke and sarcasm semantics were found to account for the largest percentage of tweets mis-categorized by both "out-of-the-box" and learning-based tools (Abbasi, Hassan, & Dhar, 2014). Furthermore, there is no standard schema for the identification of humor by human coders (c.f., Chew & Eysenbach, 2010). To examine the prevalence of political humor and its relationship to visibility, we undertook a mixed-methods analysis of tweets sent during the 2012 presidential debates.

Method

Data Collection

To explore the role of humor among connected viewers of the three presidential debates, we combined microscale participant observation with macroscale analysis of tweets sent during the broadcast.

On the evening of each debate, graduate students and faculty with subject area expertise gathered to observe and participate in live-tweeting. A large television displayed the broadcast feed while a random sample of tweets was projected onto a nearby wall. Equipped with their own devices and Twitter accounts, researchers took notes, made screenshots, and discussed unusual statements and images that they saw on television, on the projection, and in their personal streams.

From November 2011 to November 2012, we continuously collected tweets related to the election using the Gnip PowerTrack service (colloquially, the "firehose").[1] PowerTrack generated a custom, real-time stream of tweets based on a list of "filtering rules." Each rule was defined as a combination of keywords, phrases, hashtags, and usernames. Our data collection system included a mechanism for collective management of filtering rules, enabling members of the team to add or remove rules in response to emergent phenomena as well as to later reconstruct the chronology of these changes (Wang et al., 2012).

Before each debate, we developed a list of filtering rules related to the event, issues, and candidates. To capture unanticipated themes, we also continuously updated the rules during our participant observation. At the conclusion of the third debate, we conducted a post hoc round of filtering to eliminate false positives.[2] By incorporating emergent keywords, we mitigated the problems associated with predetermined lists of hashtags or keywords (Tufekci, 2014). Only 14.5% of the messages in the final corpus included *#debate* or *#debates*.

Analysis

Each debate lasted 90 min, followed by at least 30 min of televised commentary. To observe the rise and decline of Twitter activity related to the debates, we defined a six hour window of observation beginning one hour before the start of the debate. During these three periods, we observed 33,501,086 tweets matching our rules (Figure 1). For comparison, the three evenings of the debates accounted for 12.47% of all election-related tweets sent between November 2011 and November 2012. The only single day to prompt more activity on Twitter than the debates was Election Day.

The corpus included a mix of original tweets and retweets. To identify retweets, we adopted the Bruns and Stieglitz (2013) method to sort each tweet into one of the three mutually exclusive categories: *original tweet*, *edited retweet*, and *unedited retweet*. Conventionally, unedited retweets simply replicate an earlier message with a short citation string prepended. Edited retweets include additional text such as brief commentary or additional hashtags.[3] Occasionally, a tweet passed through several rounds of retweeting, significantly altering the text of the original message. In such cases, we assessed only the relationship of the two most recent tweets in the retweeting sequence. We collapsed all retweet variants into a single retweet category and classified Tweets that did not match any of the retweet criteria as original.[4] Using this method, we classified 15,740,404 (46.98%) tweets as original and 17,760,682 (53.02%) as retweets. An overwhelming majority of the retweets—16,660,358 (93.8%)—were created using the platform's built-in "retweet" feature.

The tweets and retweets in this collection were authored by 5,423,355 unique accounts, 1,863,170 (34.35%) of which were created within the previous year and 245,477 (4.53%) of which were created during the previous month. The distribution of messages sent by this population was positively skewed, ranging from 1 to 726 (Table 1). A majority of accounts (3,748,179 or 69.11%) sent at least one retweet during the observation period and 1,570,583 accounts (28.96%) *exclusively* sent retweets and no original tweets. Conversely, most accounts that sent an original tweet were never retweeted. Of the nearly four million accounts that sent an original tweet during the observation period, only 1,427,230 (37.04%) had one or more of their messages retweeted.

Table 1. Outgoing Twitter Activity per Individual Account.

Type of activity	Median	*M*	*SD*	Min	Max
Any outgoing tweet	2.0	6.09	13.67	1	726
Original tweet	1.0	2.9	6.85	0	535
Retweet of another account	1.0	3.27	9.32	0	674

Among those accounts that sent a tweet on these three evenings, most were active during the two hours of the televised debate. Across the three debates, the volume of outgoing tweets—original and retweet—surged during the first fifteen minutes of the broadcast, peaked just after the first hour, and began a steady decline approximately thirty minutes before the conclusion of the debate. Approximately 85% of all activity occurred during the debate broadcasts (Figure 2).

Identifying Humor Among Debate Tweets

Informed by participant observation during the debates, we conducted three rounds of analysis to measure the prevalence of humor in the corpus. First, we used pattern matching to filter tweets referencing highly publicized jokes. Next, we analyzed a sample of highly retweeted accounts to identify accounts that exclusively engage in humor. Finally, we analyzed the content of a large sample of individual tweets.

Debate-Defining Moments of Humor

Before the analysis, we expected to find a significant volume of Twitter activity related to three well-known gaffes. In the first debate, Mitt Romney followed a comment about defunding public broadcasting with an aside about loving Big Bird, a character from long-running television program for children. In the second, Romney described the "binders full of women" from which he planned to recruit women into his administration. In the third debate, Barack Obama responded to a comment about the state of the US Navy's armament by quipping that the military also stocked "fewer horses and bayonets" than it did a century before. Freelon and Karpf (2015) described these gaffes as "debate-defining moments" that influenced public opinion and shaped media coverage of the campaign.

In spite of their visibility in post-debate reportage, however, tweets referring to the candidates' comments about Big Bird, binders, and bayonets accounted for a small proportion of the overall activity during the debates. To identify tweets referencing these gaffes, we parsed the text of each tweet in search of "big bird," "bayonet," or "binder," terms that were unlikely to be mentioned in any other context (Figure 3).[5] In total, we found 847,119 tweets authored by 490,696 unique accounts mentioning Big Bird, binders, or bayonets. Retweeting was more common in this sample than the whole corpus. Of the total number of tweets related to these three highly publicized moments, we classified 289,551 (34.18%) as "original" and 557,568 (65.82%) as retweets. Indeed, while references to Big Bird, binders, and bayonets were relatively rare among all the original tweets (1.84%), they were approximately twice as common among the top 1% of most retweeted tweets. All the tweets and retweets related to these three gaffes were labeled "humor" in our database.

Identifying Accounts that Exclusively Tweet Humor

During our participant observation, we noted that many highly visible accounts seemed to *exclusively* send out messages of a humorous nature. Some, such as @BorowitzReport, @TheOnion, @billmaher, and @SarahKSilverman, were "verified" by Twitter and explicitly affiliated with professional comedians and satirical media organizations, while others, such as @BigBirdRomney and @RepubGrlProbs, were created during the previous 12 months and appeared to exist solely for the purpose of political humor. To identify the prevalence of these "strictly humor" accounts, we undertook a systematic content analysis of accounts.

The discourse among connected viewers of the debates was fundamentally shaped by accounts sending out *original* tweets as opposed to retweets. The visibility of these accounts was amplified, in turn, through the retweets sent by others. To facilitate a sampling strategy based on the relative visibility of individual accounts, we assigned a visibility measure to each account, constructed in two steps: First, we calculated the retweet frequency of every original tweet in the corpus. Second, we summed the retweet frequencies for each account sending one or more original tweets. We then ranked the population of accounts by retweet frequency and removed accounts with zero retweets. The resulting subpopulation included 1,427,230 accounts or 26.32% of the connected viewing population. The distribution of retweets was highly skewed. We traced 50% of all retweets back to just 3,399 accounts (0.0001% of all accounts in the corpus). Because visibility was not normally distributed, we stratified the ranked population of accounts into four quartiles (Table 2). Each quartile represented an approximately equal number of aggregate retweets.

To identify strictly humor accounts, we examined the profile page of each account in situ on twitter.com and answered the question: "Is the main purpose of this Twitter account humor, comedy, parody, or sarcasm?" This prompt was purposefully agnostic about the style or genre of humor at work. Only those accounts that appeared to *exclusively* send out messages fitting one or more of these criteria were classified as "strictly humor." The human coders assigned to this task were graduate students and faculty with subject area expertise and contextual knowledge based on their experience as participant observers during the debates. Coders were instructed to take into consideration the account's profile picture and design, the account's bio, and recent tweets. They were not asked to review the account's activity during the debates.[6]
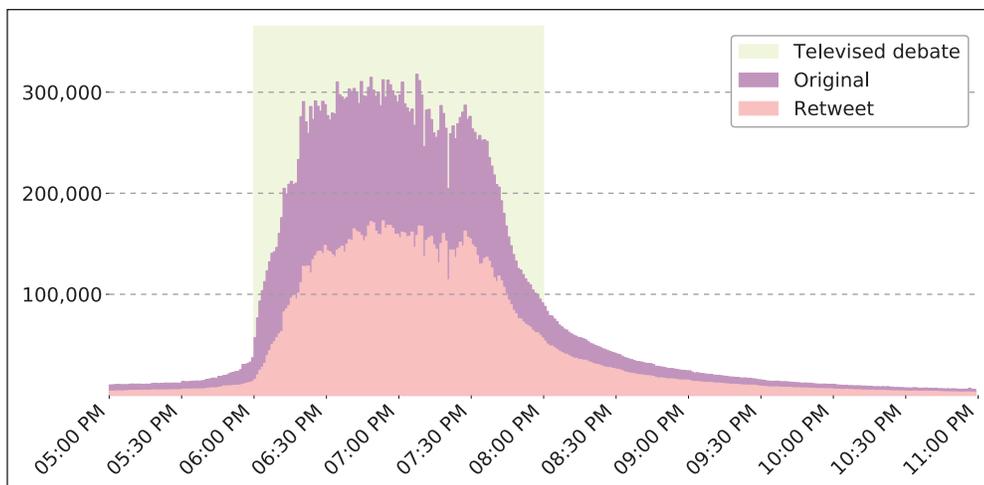
Figure 2. Live-tweeting activity summed per minute from October 3, 16, and 22, 2012. (Times listed in PDT).
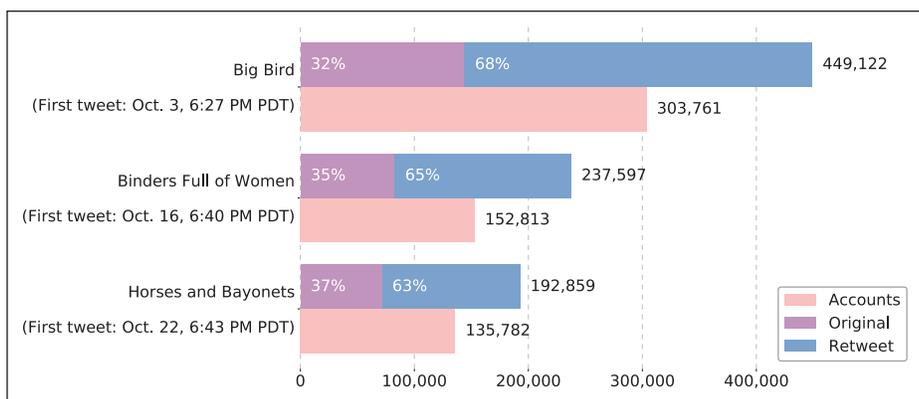

Figure 3. Highly publicized moments of humor during the three debates (not mutually exclusive).

Five coders were trained using a preliminary sample of 151 accounts (Riffe, Lacy, & Fico, 2005). The reliability sample was stratified according to the four quartiles: 14 accounts from the top-ranked quartile, 25 from the second most retweeted quartile, and 56 each from the least retweeted quartiles. Reliability scores were calculated independently for each quartile (Table 3).[7] The coders proved highly reliable only when coding accounts from the top quartile ($\alpha=0.92$), a distinction that suggested a qualitative difference between the most visible accounts and the rest of the sample.

The coders overwhelmingly agreed that the accounts from the less-visible three quartiles were not *strictly* used for humor. Of the 137 accounts in the less-retweeted groups, only 12 (or 8.76%) were identified as strictly humor by one or more coders. In contrast, 8 of the 14 (or 57.1%) accounts drawn from the top-ranked quartile were coded as strictly humor. Based on these results, we decided to analyze only the top-ranked quartile. Strictly humor accounts appeared so rarely in the lower ranked quartiles that it would be impossible to distinguish the results of human coding from random chance, a phenomenon identified by Krippendorff (2004).

After completing the reliability test, a census of the remaining highly visible accounts was distributed evenly among the coders ($N=196$). Of these, 61 accounts (31.1%) were coded as strictly humor. These accounts sent 2,619 original tweets that were, in turn, retweeted by others 1,239,552 times during the debates. In total, accounts identified as strictly humor by our coders were responsible, directly and indirectly, for 1,242,836 tweets and retweets, or 3.71% of the full corpus. All the humor accounts in the top quartile tweeted primarily in English. All the original tweets authored by strictly humor accounts and the accompanying retweets were marked as "humor" in our database.

Identifying Unambiguously Humorous Tweets

The previous analysis revealed that a small number of Twitter accounts dedicated exclusively to humor were

Table 2. Connected Viewers Retweeted Once or More Ranked and Stratified by Retweet Count.

Ranked quartiles	Unique accounts (% of the retweeted population)	Retweets of the accounts in this quartile (% of all retweets in corpus)	Outgoing Tweets	Outgoing Original (% of activity in this quartile)	Outgoing Retweet (% of activity in this quartile)
Population	1,427,230	15,772,934	18,897,898	9,838,628	9,059,270
Quartile 1	186 (0.01%)	3,947,720 (25.03%)	19,864	15,025 (75.64%)	4,839 (24.36%)
Quartile 2	3,213 (0.02%)	3,939,007 (24.97%)	224,153	153,276 (68.38%)	70,877 (31.62%)
Quartile 3	92,285 (6.47%)	3,942,981 (25%)	3,674,086	2,042,026 (55.58%)	1,632,060 (44.42%)
Quartile 4	1,331,546 (93.3%)	3,943,226 (25%)	14,979,795	7,628,301 (50.92%)	7,351,494 (49.08%)

Note. A total of 1,987,748 retweets could not be reliably traced back to an original tweet and were excluded; 1,100,324 were "manual" retweets and 887,424 were retweets of messages sent outside of the observation period.

Table 3. Intercoder Reliability for the Analysis of Accounts.

Sample	Krippendorff's α
Quartile 1	0.92
Quartile 2	0.67
Quartile 3	−0.16
Quartile 4	0.79
Total	0.78

responsible for an outsized volume of highly visible (i.e., frequently retweeted) messages. Yet, our participant observation suggested that humorous activity spread far beyond these strictly humor accounts. To estimate the prevalence of humor among accounts that were *not* strictly dedicated to humor, we carried out a quantitative content analysis of original tweets.

To prepare a population of tweets for analysis, we calculated the frequency with which each original tweet was retweeted, removed tweets with zero retweets, and ranked the remaining population. We also excluded any tweet marked as humor in the previous two rounds. The resulting population included 3,106,594 original tweets that were collectively retweeted 13,125,117 times. The distribution of retweets among original tweets was highly skewed—more than half of the tweets (57.48%) were retweeted just once—so we stratified the retweeted tweets into four quartiles representing equal shares of the overall visibility (Table 4).

In this stage of the analysis, we classified the tweets as either "humor," "non-humor," or "unsure." The coders were prompted to mark a tweet "humor" if it "appeared primarily authored for the purpose of humor" (see Appendix for the complete codebook). To avoid biases in coders' individual senses of humor, we instructed the coders to conjecture about the author's motivation in composing a tweet rather than evaluate the content alone. In practice, this allowed coders to mark a tweet as "humor" even if it did not strike them personally as funny. Coders were further instructed to label tweets conservatively, preferring to exclude rather than include edge cases. In addition to in-person training, coders were provided with 13 example tweets, 7 of which were marked as "humor" and 6 of which were marked as "non-humor." A stratified reliability sample was taken from

the four quartiles and distributed to the team of coders ($N=73$). The team underwent two rounds of reliability testing and reached a satisfactory level of intercoder reliability given the difficulty of the coding exercise ($\alpha=0.685$). Reliability scores were calculated using the *ReCal* software package (Freelon, 2010).

A stratified sample was drawn independently from each quartile of the population and combined into a single sample ($N=1,503$) to achieve a 95% confidence level. At the conclusion of the coding exercise, 327 (21.76%) tweets were identified as "humor," 1,061 (70.59%) were coded "non-humor," and 115 (7.65%) were marked "unsure" and set aside for close textual analysis (Table 5). Consistent with the analysis of accounts, the proportion of cases clearly identifiable as humor was greatest among the top-ranked highly retweeted tweets; messages sampled from lower in the rankings tended to be more difficult for coders to interpret. Using the approach described in Riffe et al. (2005, p. 109), we determined that the sampling error of proportion for the stratified sample was 0.0189 at a 95% confidence interval. Because of the strict coding instructions, these results should be interpreted as a conservative but reliable *under*-estimate of the prevalence of humor in the sample.

Findings

How Much of the Overall Connected Viewing Activity was Dedicated to Humor?

To estimate the overall proportion of connected viewing activity related to humor, we generalized from the sample of tweets coded "humor" in our database to the total corpus of tweets. To account for the skewed distribution of retweets, we stratified the total population of original tweets and calculated independent estimates for each quartile. Every original tweet was assigned to one of the four quartiles according to its retweet count (Table 6). Original tweets with zero retweets were assigned to the least visible group, quartile 4. Original tweets sent prior to, but retweeted during, the observation period were analyzed separately. After every tweet and retweet was assigned to one of the four groups, we calculated an independent estimate of humor for each group and summed the results (Table 7).

Table 4. Ranked and Stratified Population of Original Tweets.

	Total original	Total retweet	Mean retweet	Median retweet	Min retweet	Max retweet
Quartile 1	3,980	3,281,538	824.50	484	281	41,297
Quartile 2	56,139	3,281,035	58.44	37	19	281
Quartile 3	574,512	3,281,265	5.71	4	3	19
Quartile 4	2,471,963	3,281,279	1.33	1	1	3
Combined	3,106,594	13,125,117	4.22	1	1	41,297

Note. Tweets referring to Big Bird, binders, or bayonets, and tweet sent by strictly humor accounts were excluded.

Table 5. Humor Among Original Tweets Sent During the US Presidential Debates in 2012.

	Humor	Non-humor	Unsure	Total
Quartile 1	96 (27.12%)	235 (66.38%)	23 (6.5%)	354
Quartile 2	90 (23.56%)	264 (69.11%)	28 (7.33%)	382
Quartile 3	70 (18.23%)	282 (73.44%)	32 (8.33%)	384
Quartile 4	71 (18.54%)	280 (73.11%)	32 (8.36%)	383
Total	327 (21.76%)	1061 (70.59%)	115 (7.65%)	1503

Note. Quartiles are ranked 1 to 4 from most to least retweeted.

Using this stratified approach, we estimate that 6,865,493 messages, or 20.49% of the corpus, were clearly intended as humor.

Our estimates indicated that humor was more prevalent among retweets (22.22%) than original tweets (18.55%). This difference suggested, but did not prove, that tweets expressing humor were retweeted more often than non-humor tweets.

What was the Relationship Between Humor and Visibility Among Live-Tweeting Viewers?

To investigate whether humorous messages gain more visibility (i.e., receive more retweets) than non-humorous messages, we modeled a negative binomial regression to predict the number of retweets that a particular original tweet received, controlling for a number of variables, including if the tweet was categorized as strictly humor by a human coder.[8] The model allows us to infer how much the presence of humor in an original tweet impacted the resulting number of retweets. In addition to whether or not the tweet was intended to be humorous, our model included six additional independent variables that may also have contributed to the retweet count ($N=1,346$):

Number of Retweets. A dependent count variable of retweets per tweet (min=1, max=8387, $M=200.994$, $SD=538.079$).

Humor. A dichotomous variable indicating whether the tweet was coded strictly humor (yes=317, no=1,029).

Minutes from Start. The number of minutes between the start of the observation period and the time that the tweet was posted (min=1, max=315, $M=121.008$, $SD=41.977$).

Author Followers. The number of followers, logged, of the tweeting account at the time that they sent the tweet (min=1.609, max=16.871, $M=8.366$, $SD=2.868$).

URL. A dichotomous variable indicating whether the tweet contained a URL (yes=73, no=1,273).

Media. A dichotomous variable indicating whether the tweet contained media (such as an image; yes=11, no=1,335).

Hashtag. A dichotomous variable indicating whether the tweet contained one or more hashtags (yes=413, no=933).

Hashtag Count. The count of the number of hashtags in the tweet (min=0, max=7, $M=0.410$, $SD=0.738$).

The model shows that, after controlling for all other variables, tweets labeled humor were significantly more likely to be retweeted ($Exp(\beta)=2.537$, $p<.001$); humor tweets had around a 2.5 times higher log of expected counts of retweets than non-humor tweets (Table 8). Other factors that contributed positively to the number of retweets were the author's (logged) number of followers ($Exp(\beta)=1.712$, $p<.001$) and if the tweet contained media ($Exp(\beta)=1.298$, $p<.001$) or at least one hashtag ($Exp(\beta)=1.800$, $p<.001$). Multiple hashtags had a significant, lower log of expected retweets ($Exp(\beta)=0.822$, $p<.001$).

The model supports our intuition that humor contributed to the overall visibility of original tweets. Not only was a large proportion of connected viewing activity dedicated to humor, but clear expressions of humor were made more visible than other messages in the corpus.

Discussion

Retweeting is a platform-specific mechanism for the collective allocation of visibility. During the 2012 US presidential debates, a small number of accounts and messages became unusually visible after receiving an overwhelming majority of retweets. While this unequal distribution of visibility may be partially explained by contextual factors such as an account's preexisting notoriety or the mechanism by which Twitter curates search results, our analysis demonstrates the critical role of humor in the allocation of attention. Many connected viewers were afforded greater visibility specifically because of their use of humor.

Table 6. Retweeting of Original Debate-Related Messages.

	Original	Retweet	Mean retweet	Median retweet	Min retweet	Max retweet
Quartile 1	5,274	4,702,263	891.59	501	281	41,297
Quartile 2	68,470	3,942,047	57.57	36	19	280
Quartile 3	740,140	3,871,424	5.23	4	3	18
Quartile 4	14,926,520	3,257,200	0.22	0	0	2
Retweets with no original	–	1,987,748	–	–	–	–
Total	15,740,404	17,760,682	1.00	0	0	41,297

Table 7. Estimated Proportion of Humor Within the Stratified Population of Debate-Related Tweets.

	Humor proportion	Sampling error of proportion	Original tweets	Estimated humor in original tweets	Retweets	Estimated humor in retweets	Estimate of total humor
Quartile 1	27.12%	0.0236	5,274	1,430.24	4,702,263	1,275,189.97	1,276,620.20
Quartile 2	23.56%	0.0217	68,470	16,131.68	3,942,047	928,754.53	944,886.20
Quartile 3	18.23%	0.0197	740,140	134,921.35	3,871,424	705,728.33	840,649.69
Quartile 4	18.54%	0.0199	14,926,520	2,767,057.23	3,257,200	603,815.14	3,370,872.38
Retweets with no original	21.76%	0.0189	–	–	1,987,748	432,464.14	432,464.14
Total			15,740,404	2,919,540.5 (18.55%)	17,760,682	3,945,952.11 (22.22%)	6,865,492.61 (20.49%)

Table 8. Negative Binomial Regression Predicting Count of Retweets.

Variable	Exp(β)	SE
Humor (Y/N)	2.537***	0.095
Minutes from start	0.997***	0.001
(Log) author followers	1.712***	0.015
Has URL (Y/N)	0.651***	0.200
Has media (Y/N)	1.298***	0.482
Has hashtag (Y/N)	1.800***	0.159
Hashtag count	0.822***	0.100
Constant	0.855***	0.174

$N = 1,346$; *$p < .05$, **$p < .01$, ***$p < .001$.
Log likelihood: −6,831.889
Reference categories: humor (no), has_url (no), has_media (no), has_hashtag (no).

Most connected viewers participated in the allocation of visibility via retweeting. Not only did retweets account for a majority of the tweets in our collection but the majority of participants (3,748,179 or 69.11%) in our corpus sent at least one retweet and 1.6 million accounts (28.96%) exclusively sent retweets. The most prolific accounts retweeted constantly during the broadcast; 326 accounts sent at least one retweet per minute. Conversely, most of the accounts that sent at least one original tweet were never themselves retweeted. Of the nearly four million participants who sent an original tweet during the observation period, only 1,427,230 (37.04%) had one or more of their messages retweeted.

Silence and invisibility are the steady states of social media. Many legitimate forms of participation—reading, searching, direct messaging, and laughing out loud—were simply invisible to our data collection apparatus and therefore also our analysis. One hint at the size of this invisible connected viewership was a set of 1.6 million accounts (28.96% of the overall population) that sent only retweets and no original tweets during the debates. Similarly, the 2,045,350 accounts (37.71%) that sent just one original tweet were somewhat more likely than the rest of the connected viewership to send out their sole messages just before the start of the debate (i.e., signaling their otherwise silent presence). The prevalence of these quiet activities is a reminder that a truly complete population of Twitter participants would include those who read tweets without ever posting or retweeting.

In practice, the circulation of humor took many forms on Twitter during the debates. We observed both spontaneous bursts of humor in response to the live event as well as generic one-liners and accounts that exclusively trafficked in humor. While contemporaneous coverage by journalists focused on a handful of humorous moments, our analysis demonstrates that highly publicized gaffes such as Big Bird, "binders full of women," and "horses and bayonets" made up just a small proportion (2.53%) of the overall debate-related humor on Twitter. Strictly humor-only accounts, meanwhile, occupied a larger share of the corpus (3.71%). Indeed, of the top quartile of most visible Twitter accounts, we labeled nearly one in three (31.1%) as "humor-only."

Succinctly characterizing the full range of humorous themes and practices in circulation during a mass-scale live-tweeting event remains a challenge. The outsized attention paid to Big Bird, binders, and bayonets during and after the debate was not incorrect, per se, but rather reflects the mechanics of visibility on Twitter. As this research demonstrates, humorous accounts and messages are more likely to be retweeted and, therefore, afforded greater visibility on the platform.

Limitations

While our method of combining participant observation with software-assisted data collection is an improvement over studies that exclusively analyze a predetermined set of hashtags, our keyword management system may have missed some tweets sent before the addition of a particular rule (though we believe these cases to be minimal). In addition, retweet count is a relatively narrow measure of visibility, given that some tweets were reproduced in mass media news coverage. While we cannot measure the total visibility of a given message, the retweet counts we observed were accurate measures of visibility *within* the system. Accounts and tweets with high retweet counts simply have a greater chance of being seen than their peers. Finally, the definition of humor in our content analysis procedures was, out of analytical necessity, particularly conservative. A more robust approach might independently analyze multiple categories of humor defined by genre and style.

Conclusion

Humor was fundamental to the political discourse on Twitter during the US presidential debates in October 2012. This multi-method content analysis indicates that at least one in five of the tweets sent by connected viewers of the presidential debates was either itself humorous, satirical, snarky, or parodic, or else referred to the humorous discourse of other users. We also found that a higher proportion of retweets (22.22%) than original messages (18.55%) were related to humor. Regression analysis supported our intuition that the presence of humor was a significant factor in the retweet counts of original messages. Not only was humor widespread among connected viewers of the debates but humorous tweets were afforded greater visibility than other messages.

Expressions of humor in our database ranged from the spontaneous to the pre-scripted, from the carefully crafted to the crude. The overall volume of humorous tweets makes clear that any effort to use aggregated social media messages as a proxy for public opinion must contend with the challenge of interpreting humor in all of its many varieties. Whether in academic, journalistic, marketing, or advocacy contexts, it is essential that social media researchers anticipate the prevalence of humor to avoid misrepresenting the discourse unfolding online. Researchers must be aware of the limits of off-the-shelf tools for text mining and sentiment analysis. Social media systems support a wide range of themes, practices, and participants. To represent this diversity of activity with validity and care, it is essential that we adopt appropriate methods shaped by firsthand participation.

The positive relationship between humor and visibility is especially salient to individuals and organizations seeking a greater share of public attention. As scholars, journalists, and practitioners have observed, attention-seeking tactics such as "joke stealing" and "content farming" exploit the connection between humor and visibility to attract massive audiences (e.g., Bonair, 2012; D'Orazio, 2015; Gates, 2015; Knibbs, 2014; O'Neil, 2015). While these tactics were pioneered by relatively harmless spammers, they are now used for the purposes of politics, persuasion, and propaganda. Recent work by investigative journalists reveals far-right activists in the United States using humor to avoid censorship by platforms, offer plausible deniability to publishers, and gain legitimacy among otherwise moderate audiences (e.g., Bernstein, 2017; Feinberg, 2017). In short, humor has become an effective conduit for extreme, racist, and anti-social speech on social media (see also Phillips, 2015; Phillips and Milner, 2017). As a rhetorical form that provides visibility and legitimacy to unpopular ideas, humor will continue to shape the future of political communication, demanding serious methodological creativity and theoretical sophistication.

Notes

1. In contrast to the subset provided by Twitter's public Streaming API, PowerTrack promises a comprehensive stream of tweets (Driscoll & Walker, 2014).
2. Filtering rules available from the authors by request.
3. In 2012, Twitter's automated retweet function was not yet universally implemented. A number of syntactic variations persisted in the corpus (for more detailed discussion of this transition, see Bruns, Burgess, Crawford, & Shaw, 2012). Retweets created by the platform were identified easily because Gnip provided a pointer to the original tweet. To identify "manual" retweets, however, we matched the body of

the tweet against the following regular expression: /.*"@|MT @|via @|RT @([A-Za-z0-9_]+)/.

4. In this context, the term "original" refers to a formal distinction and does not refer to the originality of the ideas or expression found in the body of the tweet text.

5. Specifically, we performed a case-insensitive match against the regular expression /big.*bird|binder|bayonet/.

6. The coding took place between November 2012 and January 2013. During this period, 0.5% of accounts were deleted and 1% were suspended.

7. The reliability scores were calculated using software implemented by Lippincott and Passonneau (2008).

8. We use a negative binomial model to account for the left-skewed distribution of retweets.

References

Abbasi, A., Hassan, A., & Dhar, M. (2014). Benchmarking Twitter Sentiment Analysis Tools. In N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, … S. Piperidis (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, Iceland: European Language Resources Association (ELRA).

Attardo, S. I. (2014). *Encyclopedia of humor studies*. Thousand Oaks, CA: SAGE.

Bakliwal, A., Foster, F., van der Puil, J., O'Brien, R., Tounsi, L., & Hughes, M. (2013). Sentiment analysis of political tweets: Towards an accurate classifier. In *Proceedings of the NAACL workshop on language analysis in social media* (pp. 49-58). Atlanta, GA: Association for Computational Linguistics. Retrieved from http://doras.dcu.ie/19962/

Baym, G. (2005). The Daily Show: Discursive integration and the reinvention of political journalism. *Political Communication*, *22*, 259-276. doi:10.1080/10584600591006492

Baym, G. (2007). Representation and the politics of play: Stephen Colbert's better know a district. *Political Communication*, *24*, 359-376. doi:10.1080/10584600701641441

Bentivegna, S., & Marchetti, R. (2015). Live tweeting a political debate: The case of the "Italia bene comune." *European Journal of Communication*, *30*, 631-647. doi:10.1177/0267323115595526

Bernstein, J. (2017, October 5). Here's how Breitbart and Milo smuggled Nazi and white nationalist ideas into the mainstream [Blog]. *BuzzFeed*. Retrieved from https://www.buzzfeed.com/josephbernstein/heres-how-breitbart-and-milo-smuggled-white-nationalism

Bonair, A. (2012, December 21). *Who stole my Tweet?* Available from http://who.stolemytweet.com/

boyd, D., Golder, S., & Lotan, G. (2010, January 5-8). Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. *Proceedings of 43rd Hawaii International Conference on Systems Science (HICSS-43)*, Honolulu, HI.

Bruns, A., Burgess, J. E., Crawford, K., & Shaw, F. (2012). *#qldfloods and @QPSMedia: Crisis communication on Twitter in the 2011 South East Queensland floods (Report)*. Brisbane, Australia: ARC Centre of Excellence for Creative Industries and Innovation, Queensland University of Technology. Retrieved from http://eprints.qut.edu.au/48241/

Bruns, A., & Highfield, T. (2015). May the best tweeter win: The Twitter strategies of key campaign accounts in the 2012 U.S. election. In C. Bieber & K. Kamps (Eds.), *Die US-Präsidentschaftswahl 2012: Analysen der Politik- und Kommunikationswissenschaft* (pp. 425-442). Wiesbaden, Germany: Springer.

Bruns, A., & Stieglitz, S. (2013). Towards more systematic Twitter analysis: Metrics for tweeting activities. *International Journal of Social Research Methodology*, *16*, 91-108.

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, *14*, 1164-1180. doi:10.1177/1461444812440159

Burwell, C., & Boler, M. (2014). Rethinking media activism through fan blogging: How Stewart and Colbert fans make a difference. In M. Ratto & M. Boler (Eds.), *DIY citizenship: Critical making and social media* (pp. 115-126). Cambridge: The MIT Press.

Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak. *PLoS ONE*, *5*(11), e14118. doi:10.1371/journal.pone.0014118.

Christensen, C. (2013). Wave-riding and hashtag-jumping. *Information, Communication & Society*, *16*, 646-666. doi:10.1080/1369118X.2013.783609

Conway, B. A., Kenski, K., & Wang, D. (2013). Twitter use by presidential primary candidates during the 2012 campaign. *American Behavioral Scientist*, *57*, 1596-1610.

Crimson Hexagon's Editorial Team. (2012, October 24). Social Media Analysis Shows Romney Maintains Momentum After Final Debate. Retrieved from http://www.crimsonhexagon.com/blog/opinion/social-media-analysis-final-presidential-debate/

Dailey, K. (2012, October 18). *Big Bird and binders: Election memes explained*. Retrieved from http://www.bbc.com/news/magazine-19973271.

D'Orazio, D. (2015, July 25). Twitter is deleting stolen jokes on copyright grounds [Blog]. *The Verge*. Retrieved from https://www.theverge.com/2015/7/25/9039127/twitter-deletes-stolen-joke-dmca-takedown

Driscoll, K., & Walker, S. (2014). Working within a black box: Transparency in the collection and production of big Twitter data. *International Journal of Communication*, *8*. Retrieved from http://ijoc.org/index.php/ijoc/article/view/2171

Feinberg, A. (2017, December 13). This is the daily stormer's playbook. *Huffington Post*. Retrieved from https://www.huffingtonpost.com/entry/daily-stormer-nazi-style-guide_us_5a2ece19e4b0ce3b344492f2

Freelon, D. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science*, 5(1), 20-33. Retrieved from http://www.ijis.net/ijis5_1/ijis5_1_freelon_pre.html

Freelon, D., & Karpf, D. (2015). Of big birds and bayonets: Hybrid Twitter interactivity in the 2012 Presidential debates. *Information, Communication & Society*, *18*, 390-406. doi:10.1080/1369118X.2014.952659

Gates, S. (2015) Going viral by stealing content: Can the law cure the problem of viral content farming. *Fordham Intellectual Property, Media & Entertainment Law Journal*, *26*: 689.

Gayo-Avello, D. (2012). I wanted to predict elections with Twitter and all I got was this lousy paper: A balanced survey on election prediction using Twitter data. arXiv:1204.6441. Retrieved from http://arxiv.org/abs/*1204*.6441

Gillespie, T. (2012). Can an algorithm be wrong? *Limn*. Retrieved from http://limn.it/can-an-algorithm-be-wrong/

Highfield, T. (2015a). News via Voldemort: Parody accounts in topical discussions on Twitter. *New Media & Society*, *18*, 2028-2045. doi:10.1177/1461444815576703

Highfield, T. (2015b). Tweeted joke lifespans and appropriated punch lines: Practices around topical humor on social media. *International Journal of Communication*, *9*, 2713-2734. Retrieved from http://ijoc.org/index.php/ijoc/article/view/3611

Highfield, T., Harrington, S., & Bruns, A. (2013). Twitter as a technology for audiencing and fandom. *Information, Communication & Society*, *16*, 315-339. doi:10.1080/1369118X.2012.756053

Hodas, N. O., & Lerman, K. (2012, September 3-5). *How visibility and divided attention constrain social contagion*. Paper presented at the 2012 International Conference on Social Computing *(Privacy, Security, Risk and Trust), Amsterdam, The Netherlands*. doi:10.1109/SocialCom-PASSAT.2012.129

Holt, J., & Sanson, K. (2013). *Connected viewing: Selling, streaming, & sharing media in the digital age*. London, England: Routledge.

Holton, A. E., & Lewis, S. C. (2011). Journalists, social media, and the use of humor on Twitter. *Electronic Journal of Communication*, *21*, 1-2. Retrieved from http://www.cios.org.proxyau.wrlc.org/EJCPUBLIC/021/1/ 021121.html

Hong, S., & Nadler, D. (2012). Which candidates do the public discuss online in an election campaign? The use of social media by 2012 presidential candidates and its impact on candidate salience. *Government Information Quarterly*, *29*, 455-461. doi:10.1016/j.giq.2012.06.004

Houston, J. B., McKinney, M. S., Hawthorne, J., & Spialek, M. L. (2013). Frequency of tweeting during presidential debates: Effect on debate attitudes and knowledge. *Communication Studies*, *64*, 548-560. doi:10.1080/10510974.2013.832693

Jenkins, H., Ford, S., & Green, J. (2013). *Spreadable media: Creating value and meaning in a networked culture*. New York: New York University Press.

Jones, J. P., Baym, G., & Day, A. (2012). Mr. Stewart and Mr. Colbert go to Washington: Television satirists outside the box. *Social Research*, *79*(1), 33-60.

Katz, L. (2012, October 3). Debate makes Big Bird a big hit on Twitter. Retrieved from http://www.cnet.com/news/debate-makes-big-bird-a-big-hit-on-twitter/

Kelly, H. (2012, October 23). Funniest tweets of the final presidential debate. Retrieved from http://www.cnn.com/2012/10/23/tech/social-media/funniest-tweets-third-debate/index.html.

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, *30*, 411-433.

Lippincott, T., & Passonneau, B. (2008). *Agreement.py*. Retrieved from http://cswww.essex.ac.uk/Research/nle/arrau/alpha.html

McKinney, M. S., Houston, J. B., & Hawthorne, J. (2014). Social watching a 2012 republican presidential primary debate. *American Behavioral Scientist*, *58*, 556-573.

Mejova, Y., Srinivasan, P., & Boynton, B. (2013). GOP primary season on twitter: Popular political sentiment in social media. In Proceedings of the sixth ACM international conference on Web search and data mining (WSDM' 13) (pp. 517-526). New York, NY: ACM.

Nahon, K., & Hemsley, J. (2013). *Going viral*. Cambridge, UK: Polity Press.

Nielsen. (2012a, October 17). 65.6 million viewers watched the second presidential debate. *Newswire*. Retrieved from http://www.nielsen.com/in/en/insights/news/2012/65-6-million-viewers-watched-the-second-presidential-debate.html

Nielsen. (2012b, October 4). First presidential debate draws 67.2 million viewers. *Newswire*. Retrieved from http://www.nielsen.com/us/en/insights/news/2012/first-presidential-debate-draws-67-2-million-viewers.html

Nielsen. (2012c, October 23). Final presidential debate draws 59.2 million viewers. *Newswire*. Retrieved from http://www.nielsen.com/us/en/insights/news/2012/final-presidential-debate-draws-59-2-million-viewers.html

Nilsen, D. L. F. (1990). The social functions of political humor. *The Journal of Popular Culture*, *24*, 35-47. doi:10.1111/j.0022-3840.1990.2403_35.x

O'Neil, L. (2015, January 27). Everyone's stealing jokes online. Why doesn't anyone care? [Blog]. *The Washington Post*. Retrieved from https://www.washingtonpost.com/posteverything/wp/2015/01/27/everyones-stealing-jokes-online-why-doesnt-anyone-care/?utm_term=.76aed4637174

Papacharissi, Z., & Oliveira, M. (2012). Affective news and networked publics: The rhythms of news storytelling on #Egypt. *Journal of Communication*, *62*, 266-282. doi:10.1111/j.1460-2466.2012.01630.x

Pezzoni, F., An, J., Passarella, A., Crowcroft, J., & Conti, M. (2013). Why do I retweet? An information propagation model for microblogs. In A. Jatowt*, et al*. (Eds.), *International Conference on Social Informatics* (Paper presented at 5th International Conference, SocInfo 2013, Kyoto, Japan, 25–27 November 2013) (Vol. 8238, pp. 360-369). Cham: Springer.

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

Phillips, W., & Milner, M. R. (2017). The Ambivalent Internet: Mischief, Oddity, and Antagonism Online. Cambridge, UK: Polity.

Pittman, M., & Tefertiller, A. C. (2015). With or without you: Connected viewing and co-viewing twitter activity for traditional appointment and asynchronous broadcast television models. *First Monday*, *20*(7). doi:10.5210/fm.v20i7.5935

Riffe, D., Lacy, S., & Fico, F. (2005). *Analyzing media messages: Using quantitative content analysis in research*. New York, NY: Routledge.

Schirra, S., Sun, H., & Bentley, F. (2014). Together alone: Motivations for live-tweeting a television series. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '14)* (pp. 2441-2450). New York, NY: ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2557070

Sharp, A. (2012, August 1). A new barometer for the election. *Twitter Blog*. Retrieved from https://blog.twitter.com/2012/a-new-barometer-for-the-election

Shifman, L. (2007) Humor in the age of digital reproduction: Continuity and change in internet-based comic texts. *International Journal of Communication*, *1*, 187-209.

Shifman, L., & Blondheim, M. (2010). The medium is the joke: Online humor about and by networked computers. *New Media & Society*, *12*, 1348-1368.

Thorson, E., Hawthorne, J., Swasy, A., & McKinney, M. S. (2015). Co-Viewing, Tweeting, and Facebooking the 2012 presidential debates. *Electronic News*, *9*, 195-214. doi:10.1177/1931243115593320

Tsou, M., Yang, J., Lusher, D., Han, S., Spitzberg, B., Gawron, J. M., & An, L. (2013). Mapping social activities and concepts with social media (Twitter) and web search engines (Yahoo and Bing): A case study in 2012 US presidential election. *Cartography and Geographic Information Science*, *40*, 337-348. doi:10.1080/15230406.2013.799738

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity, and other methodological pitfalls.

In *Proceedings of the 8th international conference on weblogs and social media (ICWSM)* (pp. 505-514). Palo Alto, CA: The AAAI Press.

vanessa. (2012, October 19). Presidential debates: Big bird and binders of women. *Topsy*. Retrieved from http://topsy.thisis-thebrigade.com/blog/presidential-debates-big-bird-and-bind-ers-of-women/

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A system for real-time twitter sentiment analysis of 2012 US presidential election cycle. In *Proceedings of the ACL 2 system demonstrations* (pp. 115-120). Stroudsburg, PA: Association for Computational Linguistics. Retrieved form http://dl.acm.org/citation.cfm?id=2390470.2390490

Warner, J. (2007). Political culture jamming: The dissident humor of the daily show with Jon Stewart. *Popular Communication*, 5, 17-36.

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. Cambridge: The MIT Press.

Zhang, Y., Wells, C., Wang, S., & Rohe, K. (2017). Attention and amplification in the hybrid media system: The composition and activity of Donald Trump's Twitter following during the 2016 presidential election. *New Media & Society*, 1461444817744390. Retrieved from https://doi.org/10.1177/1461444817744390.

Author Biographies

Kevin Driscoll (PhD, University of Southern California) is an assistant professor in the Department of Media Studies at the University of Virginia. His research interests include popular culture, political communication, and the history of technology.

Alex Leavitt (PhD, Annenberg School for Communication & Journalism, University of Southern California) is a user experience researcher at Facebook Research. His research interests include politics, news, polarization, misinformation, credibility, and trust.

Kristen L. Guth (PhD, University of Southern California) is a postdoctoral scholar of Risk and Organizational Communication in the Department of Communication at the University of Kentucky. Her research interests include processes around innovation in start-ups, including change management, team collaboration, expertise negotiation, technology design, and knowledge management.

François Bar (PhD, UC Berkeley) is professor of Communication and Spatial Sciences at the Annenberg School, University of Southern California. His research interests include the social and economic impacts of information technologies, with a focus on telecommunication policy, user-driven innovation, and technology appropriation. He works at the intersection of communication technology, development, participatory design, and urban change.

Aalok Mehta is a PhD candidate at the University of Southern California. His research interests include telecommunications, Internet, cybersecurity, and media policy, and emerging technologies including artificial intelligence, autonomous vehicles, and unmanned aerial systems.

Appendix I

Codebook for the Content Analysis of Tweets

A tweet should be coded with a "1" if it appears primarily authored for the purpose of humor. Coders should label tweets conservatively, preferring to exclude rather than include edge cases.

Humor. Humor tweets come in a wide variety but there are some common types. These examples were drawn from our initial exploration of the data and should all be marked *humor*.

Pop Culture References

- *"Kanye need to interrupt Romney."*
- *"I don't like Obama or Mitt Romney. . . #ELLENDEGENERES2K12"*

Wry Observations, Often Expressing Frustration at the Whole Affair

- *"Just pin Romney and Obama in a cage together and tell them whoever survives is president. Save a lot more time and money. #debate"*
- *"I think we SHOULD outsource debate moderating jobs for 2016. Mumbai has to do a better job than this."*

Wisecracks that Criticize One or the Other Candidate but do not Pertain to the Specific Content of the Debates

- *"MR. ROMNEY How do u spell dog? Well u see, dog is an easy word, while getting a burger in AL, a family came to me and said i lost my dog. . ."*
- *"The biggest misperception about Obama is that he can't debate. Also, that he's a Kenyan Communist from space. But the first one, mainly."*

Non-humor. The presence of sarcasm or slang is not sufficient for a tweet to be coded as humor. Tweets that are not primarily motivated by humor although they are sarcastic or include slang should be coded as *non-humor*.

- *"No, no. Let's NOT. exactly how i feel about everything romney says"*
- *"These debates are pointless. They assume undecided voters can actually find their way to the polls. #debate"*

Ambiguous Cases. Ambiguous cases often rely on subtle, non-verbal cues such as tone or context. Without the help of tone or context, these tweets may be interpreted as either humor or non-humor. When faced with such messages, they should be coded as *non-humor*.

- *"If Romney becomes president, the migration of America will begin. . ."*
- *"Remember when obama had brown hair?????? #graysfordays"*