Article

Social Science Computer Review
2017, Vol. 35(1) 10-32

DOI: 10.1177/0894439315602858
journals.sagepub.com/home/ssc

Socioeconomic Status, Social Capital, and Partisan Polarity as Predictors of Political Incivility on Twitter: A Congressional District-Level Analysis

Chris J. Vargo[1] and Toby Hopp[1]

Abstract

Using 414,322 tweets drawn from 143,404 individual Twitter users located in all 435 U.S. congressional districts, this study employed big data and automated content analysis techniques to explore the degree to which socioeconomic status (SES), social capital potential (the degree to which a congressional district has the potential for interconnected citizen networks), and in-district partisan polarization were associated with incivility on Twitter during the 2012 presidential election. Broadly speaking, and with some exceptions, the results indicated that election oriented incivility on Twitter was highest in districts that had low SES indicators, low levels of social capital potential, and low levels of partisan polarity. In its sum, this study shows how large social data sets (i.e., the Census) can be combined with big data to explain social phenomena.

Keywords

incivility, big data, Twitter, social capital, 2012 general election, congressional districts, partisan polarity

Over the past two decades, the rapid proliferation of web-based, socially interactive digital platforms has provoked intense exploration of social media's democratic potential (e.g., Freelon, 2013; Gil de Zúñiga, Jung, & Valenzuela, 2012; Papacharissi, 2004; Sunstein, 2001; Valenzuela, Kim, & Gil de Zúñiga, 2012). Perhaps unsurprisingly, this exploration has yielded fairly sharp debate (e.g., Dahlberg, 2001; Davis, 2009; Stromer-Galley & Muhlberger, 2009). Perspectives framing the debate can be segmented into two camps. The first perspective holds that the accessible nature of the social Internet will encourage a relative widening of the public sphere and in so doing, "pave the road for a democratic utopia" (Papacharissi, 2004, p. 260). Conversely, the alternate perspective suggests that despite their potential, social media technologies simply reinforce existing participatory gaps

[1] University of Alabama, Tuscaloosa, AL, USA

Corresponding Author:
Chris J. Vargo, Department of Advertising and Public Relations, University of Alabama, Box 870172, Tuscaloosa, AL 35487, USA.
Email: cjvargo@ua.edu

and thereby, ''frequently induce fragmented, nonsensical, and enraged discussion'' (p. 260) that possesses inherently limited democratic value.

As social media has grown, so has the scale of online discourse. For instance, in 2012, 622 million pieces of original content were generated on Twitter alone (Goel, Anderson, Hofman, & Watts, 2015). Recent developments in the computer and social sciences have enabled wider scholarly access to very large and potentially very rich data sets containing information about millions—if not billions—of individual user interactions. Access to these so-called big data sets, coupled with increasingly sophisticated analytic tools and approaches, has enabled social scientists new avenues for the appraisal of social media's democratic potential. As a means of continuing development in this area, the present study sought to explore the degree to which off-line factors previously associated with productive political discussion and debate can be used to predict online behavior. It does so using automated content analysis, which allows the results to be scaled to the proportions necessary for evaluation of big data.

Examination of the literature suggests that off-line civic involvement has been previously associated with a host of contextual factors, including socioeconomic conditions, community heterogeneity, community in/stability, and ideologically driven partisan polarization (e.g., Costa & Kahn, 2003; Fiorina & Abrams, 2008; Kang & Kwak, 2003). Building upon this research, the current study set out to explore the relationship between contextual-level features associated with individual U.S. voting districts and aggregated (at the district level) political incivility on Twitter. Using a big data sample consisting of 414,322 tweets (drawn from 143,404 individual Twitter users located in all 435 U.S. congressional districts) relevant to the 2012 presidential election, this study was specifically interested in the degree to which socioeconomic factors, social capital potential (the degree to which a congressional district has the potential for interconnected citizen networks), and in-district political polarization were related to incivility on Twitter. Notably, and despite the obvious potential for the Internet to both widen and deepen the public sphere, relatively few studies have attempted to explore the degree to which web-based platforms such as social media sites *actually* facilitate meaningful participation throughout the citizenry. As such, this study analyzed citizen behavior using big data techniques, which offer novel means of addressing questions relevant to the relationship between social media, society, and democracy as they relate to political discussion.

Literature Review

Incivility

Drawn from the Latin word *incivilis* (which can be roughly translated as *unmannerly*, *unjust*, and, ultimately, *unbecoming of a citizen*), incivility refers to a wide array of behaviors that range from rudeness and name-calling to forcible theft and hooliganism. Civility, as the converse of incivility, is thought to be central to a well-functioning democracy. For instance, Boyd (2006) argued that adherence to the principles of civil engagement helps facilitate social interactions. Civility helps citizens ''communicate respect for others and generate habits of moral equality in the everyday of life of a democracy'' (p. 863). Specifically on the topic of civil communication, Coe, Kenski, and Rains (2014) noted that ''commitment to civil discourse—the free and respectful exchange of ideas—has been viewed as a democratic ideal from the ancient Athenian forums to the mediated political debates of modern times'' (p. 658). Grand (2014) summarized the importance of civility by simply noting, ''democracy requires democrats'' (p. 8).

Incivilities manifest in a number of behavioral forms. According to Boyd (2006), there are two inter-related connotations of civility. The first of these connotations, referred to by Boyd as *formal civility*, speaks to the manner and tone in which interactions are carried out in everyday life. The second connation, understood as *substantive civility*, denotes a sense of membership in the

surrounding social and political community. These dimensions are inherently related to one another, as the former (formal civility) can be understood as exerting a direct influence on the latter (substantive civility). Specifically, Boyd notes:

> civility is a kind of 'adverbial' restraint on the civic language we speak with one another. In the same way that one is enjoined to speak politely, modestly or temperately, the adverbial condition of civility modifies and qualifies conduct without specifying its content. (p. 864)
>
> Modern society is perforated by concerns that the age of civility (if indeed such a thing ever existed) is in its twilight. Forni (2011), for instance, remarked: In today's America, incivility is on prominent display: in the schools, where bullying is pervasive; in the workplace, where an increasing number are more stressed out by co-workers than their jobs; on the roads, where road rage maims and kills; in politics, where strident intolerance takes the place of earnest dialogue; and on the Web, where many check their inhibitions at the digital door. (par. 1)

Partially in response to such concerns, communication scholars have increasingly begun to investigate the issue of incivility, particularly as it relates to online political discussion (e.g., Coe et al., 2014; Hmielowski, Hutchens, & Cicchirillo, 2014; Papacharissi, 2004; Rowe, 2015; Santana, 2014). Operating under the assumption that computer-mediated civil political discourse is a meaningful component of 21st-century political involvement, these studies have focused on the individual user and addressed topics such as the role of usage patterns relative to behavior (e.g., Hmielowski et al., 2014; Papacharissi, 2004), the relationship between media content/structure and discussion tone (e.g., Borah, 2014; Coe et al., 2014), and the effects of uncivil discussion on issue perceptions (Anderson, Brossard, Scheufele, Xenos, & Ladwig, 2014). Focus on individual-level attributes through experiments and surveys helps build a comprehensive understanding on the social effects of incivility. It generally does not, however, address incivility as a phenomenon that is rooted in surrounding social structures. Moreover, these studies tend to address citizen behavior prospectively rather than when it actually occurs.

Accordingly, the goal of this work was to explore the degree to which contextual factors associated with social life are associated with aggregated patterns of computer-mediated political incivility. In so doing, this study attempted to empirically investigate the degree to which factors commonly associated with substantive incivility (i.e., membership in the surrounding social and political community) can also be used to describe patterns of formal incivility (i.e., the manner and tone in which social interactions are carried out on Twitter).

Social Media and Civic Participation

Previous research has explored civic participation in the context of social media use (e.g., Gil de Zúñiga et al. 2012; Lutz, Hoffman, & Meckel, 2014; Valenzuela, Park, & Kee, 2009). Research in this area has found that social media can promote civic involvement. In so doing, it can increase public awareness and optimism about social projects (Marinov & Schimmelfennig, 2015). In a broad sense, digital media use has bolstered the amount of political talk for those with lower levels of political interest (Bimber, Cunill, Copeland, & Gibson, 2015). Those who are generally unmotivated to talk about politics do seem to talk more when they use digital media to consume news. However, some debate exists as to where that "talk" takes place and the degree to which it is beneficial to society. In a study of newspaper comments during the 2012 general election, Coffey, Kohler, and Granger (2015) found that the majority of online discussion was uncivil. The commentators described the language as "shocking" and "disappointing" and went on to say that the discussion "contained not just insults about the candidates but designations of anyone who supported a candidate" (p. 262). The authors did find thoughtful arguments among the comments but were "somewhat pessimistic" as the degree to which social media can facilitate meaningful political discussion.

Indeed, examination of the literature indicates that social media's potential for meaningful civic engagement is neither uniform nor unconditional (Vissers & Stolle, 2014). For instance, some have noted that online participation varies greatly from person to person and from topic to topic (Lutz et al., 2014). Yet others have observed that the quality of online political participation is also influenced by number of contextual variables relating to one's social surroundings (Coffey et al., 2015). By specifically investigating how social, economic, and political conditions influence aggregate levels of political civility, this study addresses the latter concern. In so doing, our hope is to add additional clarity to the rapidly growing (Lutz et al., 2014) body of literature on political participation and social media.

Socioeconomic Status (SES)

SES has long been understood as a key predictor of social cohesion and civic behavior (Boardman & Robert, 2000). As summarized by Oliver (1999), prior research has consistently indicated that SES factors and civic participation are linked to one another. Specifically, the "underlying theme throughout these works is that people who are surrounded by more participators (i.e., the educated and affluent) feel more social pressure and are given more opportunities to participate themselves" (p. 190). A number of theoretical perspectives support the notion that the surrounding socioeconomic climate influences social and political behavior. *Social disorganization theory* (Shaw & McKay, 1942), for instance, holds that socioeconomic conditions play an instrumental role in a given community's ability to come together around common goals and, in so doing, cooperatively address chronic problems afflicting the community (Aiyer, Zimmerman, Morrel-Samuels, & Reischl, 2015). Specifically, community factors such as high poverty and unemployment rates limit a community's ability to control and organize behavior, subsequently resulting in diminished levels of social trust. Relatedly, classic theoretical approaches centered on *political socialization* (e.g., Hyman, 1959; Settle, Bond, & Levitt, 2011) hold that more affluent communities allow parents to devote more resources to child-rearing, particularly as it relates to shaping worldview constructions, political knowledge sets, and productive political/civic behaviors (Jennings & Niemi, 1974; McIntosh, Hart, & Youniss, 2007). Comparatively affluent areas also are more likely to possess primary and secondary school systems that have more experienced teachers, offer increased opportunities for extracurricular participation, and feature curricular approaches that include civics education (e.g., McFarland & Thomas, 2006; Settle et al., 2011; Wegner, 1991). Still other models, such as those describing chronic stress imposed by poverty (e.g., Baum, Garofalo, & Yali, 1999; Steptoe & Feldman, 2001), hold that low SES negatively affects physical and social well-being, which in in turn is associated with feelings of powerlessness and poorer coping skills. Finally, as pointed out by Brady, Verba, and Schlozman (1995), a lack of available resources may inhibit the socioeconomically disadvantaged from perceiving themselves as stakeholders in the surrounding democratic society and thus such citizens may be comparatively more inclined to engage in uncivil behavior.

Relative to communication behaviors, previous research has found a direct link between SES factors and uncivil communication practices. In a study of 400 undergraduate students, Glascock (2014) found that a composite variable describing (perceived) neighborhood quality was significantly and negatively linked to the propensity to use verbally aggressive/abusive language. In Brady et al.'s (1995) explication of their *political resource model*, the authors found that resource paucity was associated with a relative lack of the communication and organizational skills necessary to facilitate productive political discussion. Relatedly, Dahlgren's notion of *civic culture* (2000, 2003, 2005) suggests that the facets of the sociocultural world that constitute preconditions for democratic participation influence all forms of civic and political practice, including those practices related to citizen engagement in the online public sphere. Finally, previous research on structural influences affecting online political discourse has found sizeable associations between SES indicators and

political knowledge (e.g., Kwak, Williams, Wang, & Lee, 2005; McLeod & Perse, 1994). Political knowledge, in turn, has been found to influence political discussion in terms of both frequency and quality (e.g., Cappella, Price, & Nir, 2002; Evelend & Hively, 2009). Together, these studies suggest that online behavior is inextricably linked to citizens' lived experiences and personal resources (Dahlgren, 2005).

Building on the foregoing, SES factors were included in the current investigation. Such inclusion is vital because the resources necessary for civil political discourse tend to be biased toward in those in higher SES categories, who are more likely to have skills such as "a good vocabulary, the ability to communicate in English, a sense of personal efficacy, the ability to write or speak well, and the cognitive wherewithal to draw on previously existing political knowledge" (Brundige & Rice, 2010, p. 147). Furthermore, if it is indeed the case that citizens who exist in low SES contexts feel comparatively excluded from the social and political mainstream (Brady et al., 1995), it logically follows that they would be more likely to express sentiment that deviates from what is broadly considered civil. Interpreted relative to the objectives of the current study, then, the above literature suggests that citizens in low SES districts may be more likely to engage in uncivil discussion online.

Notably, SES is generally measured using a diverse array of indicators representing various facets of social life. While there is not universal agreement on the most precise indicators of SES, researchers have typically used measures such as annual household income, education level, and occupational status (Conger & Donnellan, 2007; Glascock, 2014). Following this research, we employed four measures of SES. First, we suggested that mean household income would be negatively related to political incivility on Twitter (e.g., Baum et al., 1999; McIntosh et al., 2007; Shaw & McKay, 1942). Second, we predicted that the education level of the district would be negatively associated with incivility (e.g., Hyman, 1959; Wegner, 1991). This assumption was rooted in the perception that college-educated adults are more likely to possess knowledge and skills relevant to social and political process (and their importance) and thus be more likely to meaningfully and productively engage in civic activities. Third, we suggested that districts with high unemployment rates would be associated with increased levels of political incivility as residents may feel increasingly distressed by their surrounding economic conditions (e.g., Boardman & Robert, 2000). Fourth, building on an emerging body of literature that suggests that access to health care is a reliable indicator of SES (e.g., Baum et al., 1999; Heck & Parker, 2002), we predicted that the percentage of respondents with health insurance would be negatively related to political incivility on Twitter. These hypotheses are presented below:

Hypothesis 1a: Political incivility on Twitter will be negatively related to district income level.

Hypothesis 1b: Political incivility on Twitter will be negatively related to district education level.

Hypothesis 1c: Political incivility on Twitter will be positively related to district unemployment rate.

Hypothesis 1d: Political incivility on Twitter will be negatively related to percentage of citizens in each district with health insurance.

Social Capital

The theory of social capital has a long, profound, and varied intellectual history that includes, but is not limited to, thinkers such as de Tocqueville, Durkheim, and Bourdieu. At its core, the theory holds that associations in networks of citizens help "to 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" (Newton, 2001, p. 201). The theory of social capital presupposes that a civil and

engaged society is based, to no small degree, on individual citizens' access to "network ties of goodwill, mutual support, shared language, shared norms, social trust, and a sense of mutual obligation that people can derive value from" (Huysman & Wulf, 2004, p. 1; cf. Ellison, Steinfeld, & Lampe, 2006, p. 7). Broadly speaking, social capital is an elastic concept that is commonly thought of as both a *cause* and an *effect* (Ellison, Steinfeld, & Lampe, 2007).

Social capital can be compartmentalized into two distinct types: *bridging* capital and *bonding* capital. As defined by Gittell and Vidal (1998), bridging social capital refers to ties, primarily weak in nature, that bring together groups of people previously unfamiliar with one another. Conversely, bonded social capital refers to the type of social capital, generally predicated upon the existence of strong ties, "that brings closer together people who already know each other" (p. 15; cf. Yuan & Gay, 2006, p. 1067). Despite the semantic association with "weak" and "strong" ties, both types of social capital play an instrumental role in the maintenance of behaviors (i.e., open exchange of ideas, task-related research exchange, new knowledge generation, and distribution of social support) necessary for a civil and productive society (Granovetter, 1973; Yuan & Gay, 2006). Specifically, bonding social capital facilitates in-group cohesion, which facilitates interpersonal trust and knowledge sharing (Krackhardt, 1992). Alternately, bridging capital encourages connections between otherwise disconnected networks of people and helps both "facilitate the exchange of information between distinct groups" and expedite "the flow of ideas among groups" (Kavanaugh, Reese, Carroll, & Rosson, 2005).

Prior research has measured social capital on both individual and aggregate levels. On the individual level, prior research has operationalized social capital in terms of the quantity and frequency respondents engage in pro-social civic behaviors (e.g., Blanchard & Horan, 1998; Onyx & Bullen, 2000), political and organizational participation (e.g., Wellman, Haase, Witte, & Hampton, 2001), degree of interpersonal trust (e.g., Lee & Lee, 2010; Shah, 1998), and life contentment (e.g., Shah, Kwak, & Holbert, 2001). On the aggregate level, researchers have conceptualized high levels of community-based social capital as manifest in aggregated composites of individual-level indicators of social capital (e.g., Hendryx, Ahern, Lovrich, & McCurdy, 2002), fraction of the eligible participation who voted in recent elections (e.g., Chamlin & Cochran, 1995; Rosenfeld, Messner, & Baumer, 2001), and structural opportunities for community action (Israel, Beaulieu, & Hartless, 2001).

Social Capital Potential

As the current methodological approach limited our ability to directly measure social capital, this study was instead primarily concerned with the *potential* for social capital formation. Social capital is, centrally speaking, concerned with individual citizen's access to networks built around social trust and norms of reciprocity. It can thus be reasoned that the presence (or absence) of certain social factors can serve as a conduit for the formation and maintenance of social capital (in both *bridged* and *bonded* forms). At least two factors, existing on the contextual level, should encourage the formation of social capital. First, prior research has shown that racial heterogeneity hampers relationship building and, thus, is negatively associated with social capital formation (e.g., Costa & Kahn, 2003; Yuan & Gay, 2006). Second, research has shown that neighborhood stability allows citizens to form and maintain the networks necessary for the production of social capital, which therefore acts as a vanguard against uncivil behavior. Moreover, we reasoned that familiarity with the broader cultural and social norms that govern local social behavior should encourage social capital formation and lessen the prevalence of uncivil online communication behavior. Both of these factors are reviewed in the paragraphs below.[1]

Racial diversity simultaneously offers benefits and imposes costs on communities. Gurin, Nagda, and Lopez (2004), for instance, found that campus-wide racial heterogeneity was associated with a number of positive outcomes related to on-campus democratic engagement, while Costa and Kahn

(2003) argued that racial diversity helps communities retain "epicurean variety" and resist the so-called "urban sprawl" (p. 103). Perhaps unfortunately, however, racial heterogeneity has been consistently associated with low levels of community-wide social capital (e.g., Costa & Kahn, 2003; Rupasingha, Goetz, & Freshwater, 2006). According to Yuan and Gay (2006), the tendency for racial diversity to be associated with low levels of social capital formation can be explained through a number of inter-related sociopsychological theories, including *self-categorization theory* (e.g., Turner, 1987), the *similarity–attraction hypothesis* (e.g., Byrne, 1971), and the *theory of homophily* (McPherson & Smith-Lovin, 1987). Put most simply, these theories hold that "similarity breeds connections" (McPherson, Smith-Lovin, & Cook, 2001, p. 415) and that society has embedded within it a "natural aversion to heterogeneity" (Alesina & La Ferrara, 2000, p. 225). Building upon this logic, the current study suggested that communities with high levels of racial fragmentation have lower potential for social capital production, as citizens are likely to exercise natural preferences for homogenization by restricting their social interactions to densely similar subgroups. We further suggested that such diminished social capital potential would be associated with increased levels of incivility, resulting in the following hypothesis:

Hypothesis 2a: High levels of racial heterogeneity will be positively associated with political incivility on Twitter.

Scholars of community integration and attachment have long noted the negative relationship between residential mobility and social capital production (Kang & Kwak, 2003). Referred to as the *systems model of community attachment* (Kasarda & Janowitz, 1974; cf. from Kang & Kwak, 2003), there exists fairly robust empirical support for the notion that length of residence is positively correlated with interest in local/community affairs (e.g., Oropesa, 1992), participation in local government/affairs (e.g., Kasarda & Janowitz, 1974), and the maintenance of local friendships (e.g., Liu, Ryan, Aurbach, & Besser, 1998). As noted by Kan (2007), "rapid inflows and outflows of residents in a neighborhood lead to neighborhood instability" (p. 437), which, in turn, results in social capital deficits. Broadly speaking, the relationship between social capital and residential stability is multidirectional. Residential outflow results in social capital reduction, while social capital dearths are, themselves, an impetus for neighborhood abandonment. This study predicted that districts featuring higher levels of residential stability would also be more likely to feature high levels of social capital and, therefore, be associated with diminished levels of incivility.

Hypothesis 2b: High levels of neighborhood stability will be negatively related to political incivility on Twitter.

Partisan Polarity

The last two decades have seen increased public and scholarly concern over partisan polarization in the United States (e.g., Brady & Han, 2006; Campbell, Rockman, & Rudalevige, 2008; DiMaggio, Evans, & Bryson, 1996; Fiorina & Abrams, 2008). Much of this research has focused on partisan polarization of elites (e.g., Hetherington, 2001). A smaller but substantive body of research has explored *mass polarization*. Often this research has targeted outcomes such as public opinion and voting behaviors (e.g., Klinkner & Hapanowicz, 2005; Nunn & Evans, 2006). Perhaps echoing the larger debate on partisan polarization, there does not seem to be an identifiable consensus on if and how contextual polarization influences civic and political behaviors (Fiorina & Abrams, 2008).

Careful consideration of the available literature does, however, provide some useful clues relative to the present research objectives. For the purposes of the present work, we center our discussion on the notion that that hotly contested districts (i.e., *swing* or *battleground* districts) are subject to

increased attention from both presidential candidates (Goldstein & Freedman, 2010) and, as such "represent more heterogeneous environments, where people are more likely to encounter counter-attitudinal messages" (Wolak, 2006, p. 354) both in the form of interpersonal interactions and political advertisements. Citizens in so-called battleground districts encounter a greater number of election-relevant messages and, therein, a greater number of messages that may be counter-attitudinal in nature. Indeed, Coffey et al.'s (2015) analysis of newspaper comments in battleground states found evidence that campaign environments tend to intensify the levels of online incivility. Other research has suggested the existence of a *stimulation effect* whereby conflict-inducing political advertisements (i.e., negative and attack ads) mobilize voter involvement (e.g., Goldstein & Freedman, 2002; Lau, Sigelman, Heldman, & Babbit, 1999), particularly as it relates to attentional resources (Cheng & Riffe, 2008). Specific to social media, Settle and colleagues (2014) found that citizens in battleground states are more likely to post election-relevant Facebook content than those in non-competitive states. Given that prior scholarship has suggested a relationship between the tone used in political advertising and the tone used by citizens (e.g., Cho, 2013) and that political advertising has become increasingly negative in nature (Farnam, 2012; Rubenstein, 2014), it stands to reason that citizens in battleground districts might be prone to engage in comparatively higher levels of uncivil discourse.

Taken as whole, the above research suggests that hotly contested political districts feature an increased emphasis on the on-hand election. In turn, this stimulates citizen attention on the election (Cheng & Riffe, 2008; Settle et al., 2014). Given the notions that the tone of major political campaigns has turned increasingly negative in nature (Rubenstein, 2014) and that citizen discourse is thought to be sensitive to the tone of the political messages that dominate the information environment (Cho, 2013), we reason that noncompetitive districts will feature comparatively lower levels of Twitter-based incivility:

Hypothesis 3: District partisan polarity will be negatively associated with political incivility on Twitter.

Method

Retrieving the Data From Twitter

Version 1.0 of Twitter's application programming interface (API) was called to download relevant Tweets during the 2012 presidential election period. Specifically, the streaming API call was used to download public messages from Twitter that mentioned the terms "Obama" or "Romney." The data were collected at a large scale, with a total of 70 million tweets being collected in all. The collection started on August 1, 2012 and ended on the Election Day, November 6. In all, 465,582 Tweets were collected with GPS coordinates. Here users disclosed their location at the time the tweet was sent.

Version 3 of the Sunlight Foundation's Congress API was used to resolve GPS coordinates (latitude and longitude) to correct congressional district in which the message came from (Sunlight Congress API, 2014). In total, 414,322 messages from 143,404 Twitter users were successfully resolved to congressional district.

Several concerns guided our decision to conduct the current analyses at the district level. First, given our interest in partisan polarity, the district level was the most granular level at which we could obtain accurate estimates of voting behavior. Second, we encountered difficulties resolving user-provided GPS coordinates to more precise/smaller units of analysis. Third, we had concerns that resolving to the county or neighborhood level could result in starkly unequal cluster sizes, especially

given that Twitter is more heavily used in urban areas and among younger users (Duggan, Ellison, Lampe, Lenhart, & Madden, 2015).

The Use of Geotagged Messages From Twitter

For a Tweet to be included in our data set, a Twitter user must have allowed message geotagging. Twitter prompts users to make a decision on this, typically once. Other studies have survived this same limitation and successfully predicted phenomena such as infectious disease transmission and breaking news stories (e.g., Sadilek, Kautz, & Silenzio, 2012). In fact, researchers have noted key advantages of using geotagged data as it pertains to assessing public sentiment on an issue or event. For example, the vast majority of "spam bots" that plague Twitter do not "opt-in" to geotagging of their tweets (Guo & Chen, 2014). Instead, they only provide a minimum amount of metadata alongside their tweet.

Pivoting to User

The data was then pivoted by user ID. For each user in the data set, an average location was derived. If a user sent messages from more than one district, the district in which the majority of messages originated from was treated as that user's home district. For each user, an average incivility score was calculated. This was done to limit any one user in a district from inflating district-level results. Finally, all users for each district were then averaged, and average incivility scores were created for each district.

Census Data

Data from the 2013 American Community Survey (1-year estimates) were used to obtain and derive SES and social capital data points for each congressional district (United States Census Bureau, 2014). These data were collected in 2012, during the time of the election period.

Measures

Incivility. Incivility was the criterion measure in the current study. Given the size of the data, the corpus could not be manually annotated. The researchers employed computer-assisted content analysis to derive incivility scores using Python. A script was written to process the wordlists and to develop incivility scores for the unit of analysis (individual tweets). It processed the list and detected the presence of words in the lists in the tweets.

Not unlike inter-coder agreement with manual content analysis, "algorithms and dictionaries must often be repeatedly revised and tweaked to improve their performance" (Zamith & Lewis, 2015, p. 4). This iterative process only concludes when the analysis yields a satisfactory level of construct validity. This is assessed when the researcher evaluates the algorithmic coder against the same coding decisions humans. In this case, two coders both must agree with each other to establish "gold standard" data. Then, the computer and the humans must agree at an acceptable level. Here, given the blatant nature of incivility in Twitter messages, human-to-human agreement of the two coders produced no errors.

Finally, the researchers draw attention to the point that once a computer has been verified to be valid, it will also be reliable, as computers are persistent and consistent and not prone to human error. Thus, calculating reliability is not necessary (Riffe, Fico, & Lacy, 2014; Zamith & Lewis, 2015).

Previous research by Coe, Kenski, and Rains (2014) and Papacharissi (2004) has proposed that incivility manifests in one or more of the following forms: name-calling, aspersion, lying, vulgarity, and pejorative speech. Similarly, Santana (2014), defined incivility as having nine key components:

(1) name-calling; (2) threats; (3) vulgarities; (4) abusive or foul language; (5) xenophobia; (6) hateful language, epithets, or slurs; (7) racist or bigoted sentiments; (8) disparaging comments on the basis of race/ethnicity; and (9) use of stereotypes.

In the current study, we adopted Santana's definition (with some modifications) of incivility as it broadly encompassed the conceptual approaches used by Papacharissi (2004) and Coe et al. (2014) while also including constructs potentially relevant to the 2012 presidential election. An initial, manual scan of 800 tweets revealed that eight of the nine concepts were present in tweets. Xenophobia was not present in this scan and did not appear to be relevant to the election. The researchers then leveraged two wordlists that are regularly used to detect concepts threats; vulgarities; abusive or foul language; hateful language, epithets, or slurs; racist or bigoted sentiments; and disparaging comments on the basis of race/ethnicity on short informal text on the web. Google's "bad word list" and ClueBot's "insult list" have been used by dozens of machine learning data scientists seeking to automatically detect insults in social commentary (Mueller, 2012). "Strong words" have been shown by others to be the most effective search terms. For instance, prior research has found that terms such as "idiot" and "moron" were strong predictors of incivility (Coffey et al., 2015). We could not identify an existing wordlist for name-calling; thus, this construct was developed manually.

Starting with over 600 words linked to incivility, the researchers proceeded with four rounds of manual content analysis. Eight hundred random tweets were pulled from the data set. Each tweet was scored for incivility using the initial wordlists. A Python script took stemmed words and looked for a combination of windowed and unwindowed matches. The score was calculated by adding the number of words that appeared in that tweet. In addition, when the matching word was in uppercase, or when the message contained an exclamation point, tweets were boosted by an additional point, as most affective boosting tools similarly do (Thelwall, Buckley, Paltoglou, Cai, & Kappas, 2010). The researchers read each message and (1) verified that the words flagged were indeed incivil; (2) scanned the message to see if any additional words in the tweet were uncivil, but undetected; and (3) verified that the overall score was correct. In cases where the concept of name-calling was incorrect, the researchers altered the key words (i.e., adding a word window around "hell" so "hello" was not detected). When the concept of threats was incorrect, the researchers added the appropriate key word. In all cases, the concept of vulgarities was satisfied and verified that the Python code was programmed correctly. This process was performed four times, and each time the key word list was adjusted. After each iteration, the incivility measure improved across the concept of name-calling and threats. The final percent agreement with the manual annotations was 98.5% for the concept of name-calling and 98% for the concept of threads. In all, the researchers were left with a list of 650 key words that were indicative of uncivil discourse. While this wordlist is powerful on the corpus at hand, its validity is likely limited to the context of the 2012 election, specifically short informal text that mentioned either Obama or Romney.

As a result of these procedures, a single index was created. Each user was assigned an average incivility score based upon their normal pattern of interaction. These scores ranged from 0 (*completely civil*) to 23.00 (*highly incivil*). To establish criterion validity, we next explored the degree to which the presently derived construct correlated with user-averaged arousal and sentiment. Arousal and sentiment are commonly identified constructs, and thus we were able to use existing, previously validated classification schemes. As such, sentiment and arousal were coded with the same parameters recommend by Vargo (2014). Our expectation was that incivility should be negatively associated with sentiment and positively correlated with arousal. In both cases, bivariate correlations were in the direction expected (sentiment, $r = -.22$, $p < .01$ and arousal, $r = .08$, $p < .01$).[2]

Finally, using the mean individual scores assigned to each user, we then calculated a mean incivility score for each congressional district. Figure 1 provides a visual representation of the average incivility scores associated with each U.S. congressional district.[3]

SES indicators. SES factors of interest in the current study included mean annual household income in each congressional district, the percent of individuals in each district with a bachelor's degree or higher, district unemployment rate (represented as a percentage of the total workforce), and percent of individuals in the district with health insurance.

Social capital potential indicators. Based upon the literature, two factors relevant to each district's potential for social capital were employed: racial heterogeneity and residential tenure. Residential tenure was conceptualized as the percent of the population with more than 12 years of tenure in their current residence. Racial heterogeneity was calculated at the district level using the index previously employed by Costa and Kahn (2003). The measure was calculated as:

$$\text{Heterogeneity} = 1 - \sum_{k} s_{ki^2},$$

where k represents the number of racial categories recorded in the Census data set (White, Black, American Indian, Asian, more than one race, and other) and s_{ki} represents the share of each racial category in voting district i. Higher scores on the index represented greater in-district heterogeneity.

Notably, the structure of the Census data framed the operationalization of some of these variables. For instance, the response categories associated with residence tenure were "Moved in 2010 or later," "Moved in 2000 to 2009," "Moved in 1990 to 1999," and so on. Judging that the formation of meaningful social capital with one's neighbors would be an involved process that could, conceivably speaking, take more than 4 years, we used a somewhat conservative approach to formation of the residence tenure variable.

District polarity. District polarity was assessed using a modified version of Cook's Partisan Voting Index (PVI; Wasserman, 2013). The PVI is based upon the district's voting behavior in the previous two presidential elections (2008 and 2012). Although the Twitter data were harvested before the 2012 election, we nonetheless used the results of the 2012 election as part of this measure because, we reasoned, it provided an accurate depiction of the polarization climate within the district at the time of data collection. Without modification, the PVI describes the direction (Republican or Democrat) and magnitude of partisan polarity in each district relative to the nation as a whole. The PVI was calculated as:

$$\text{PVI} = \left(\frac{pA + pB}{2}\right) - \left(\frac{PA + PB}{2}\right),$$

where pA represents the percentage of the two-way presidential vote that Obama received in the district in 2008, pB represents the percentage of the two-way vote that Obama received in the district in 2012, PA represents the percentage of the two-way national vote that Obama received in 2008, and PB represents the percentage of the two-way national vote that Obama received in 2012. A large positive value suggested that the district was polarized in favor of the Democratic Party, while a large negative value suggested polarization in favor of the Republican Party. Values near 0 were indicative of a relative lack of partisan polarity. As the current study was interested in assessing partisanship magnitude *irrespective* of party affiliation, we took the absolute value of each district's PVI value.

Control variables. All hypotheses tests controlled for a number of potentially confounding factors. First, as there are substantial differences in district population sizes across the United States, we controlled for number of inhabitants in each district. The median age of each district was controlled for, as Twitter use is heavier among younger users. Similarly, the percent of urban inhabitants was

Figure 1. Visual representation of Twitter incivility in each U.S. congressional district. An interactive version of this map can be viewed at chrisjvargo.com/incivility.

Table 1. Descriptive Statistics for Items Included in Analyses.

Indicator	Range	*M*	*SD*	Skewness	Kurtosis
District population	523,170–1,015,165	725,246.87	36,620.59	0.48	15.09
Median age	27.70–52.30	37.82	3.65	0.20	0.76
% Urban	23.51–100.00	80.72	19.01	−0.75	−0.53
Number of users	82–1,436	329.66	158.13	2.11	8.16
Average tweets per user	1.53–7.08	2.84	0.76	1.96	5.65
Racial heterogeneity	0.06–0.75	0.38	0.16	0.11	−1.00
Residence tenure (%)	13.84–50.70	31.94	6.30	0.00	−0.31
Household income (US$)	36,896.00–155,093.00	73,564.20	19,989.20	1.36	2.17
Unemployment rate (%)	2.60–19.10	8.54	2.48	0.83	0.12
% With bachelor's degree/higher	8.30–69.16	29.26	10.18	0.87	0.79
% With health-care coverage	61.53–97.39	85.48	5.71	0.12	1.52
Partisan polarity	0.02–44.00	12.50	8.96	0.91	0.42
Average incivility	0.09–0.53	0.28	0.08	0.37	0.16

controlled for, as Twitter users tend to reside in urban areas (Duggan et al., 2015). Finally, the effects of number of users and average tweets per user in each district were accounted for as a means of ensuring that the incivility measure was not upwardly biased by districts that housed a comparatively small number of infrequent but highly incivil Twitter users.

Descriptive statistics for each measure are provided in Table 1. Zero-order correlations for all of the variables included in the hypotheses tests are presented in Table 2.

Results

Hypotheses H1–H3 were tested using hierarchical ordinary least square (OLS) regression. Classes of independent variables (i.e., social economic status indicators, social capital potential indicators, and district-wide partisan polarity) were examined both individually and in conjunction with one another.

Hypotheses 1a–d were broadly concerned with the relationship between district SES and Twitter incivility. The data indicated that the mean household income was a positive predictor of incivility on Twitter ($\beta = .25$, $p < .01$) in the model including only the SES indicators; however, this relationship disappeared in the all-entry model. Given that we predicted a negative relationship between annual household income and incivility on Twitter, no support for Hypothesis 1a was observed.

Percent of users in each district with a bachelor's degree or higher was a negative and significant predictor of incivility on Twitter in both the SES-only and all-variable models ($\beta = -.62$, $p < .001$ and $\beta = -.50$, $p < .001$, respectively), thus Hypothesis 1b was supported.

Unemployment rate was positively and significantly related to incivility in the SES-only model ($\beta = .20$, $p < .001$). However, this relationship disappeared in the all-variable model. Thus, no support was found for Hypothesis 1c.

Finally, the percent of the district with health-care insurance was not related to Twitter incivility in either the SES indicator model or in the all-entry models. Accordingly, Hypothesis 1d was not supported.

Hypothesis 2a suggested that ethnic heterogeneity would be positively associated with uncivil communication on Twitter. As seen in Table 3, ethnic heterogeneity was positively and significantly related to incivility on Twitter when examined both individually ($\beta = .32$, $p < .001$) and after controlling for the effects of all other independent variables ($\beta = .28$, $p < .001$); therefore, Hypothesis 2a was supported.

Table 2. Zero-Order Correlations Between Items.

	1	2	3	4	5	6	7	8	9	10	11	12	13
District population (1)		−.17***	.14**	.12*	.00	.13**	.15**	−.10*	−0.06	.15**	−.26***	.04	−.11*
Median age (2)			−.34***	−.16**	.11*	.16**	.14**	−.21***	.50***	−.50***	.48***	−.37***	−.20***
% Urban (3)				.25***	.01	.42***	.45***	.17***	−.18***	.59***	−.39***	.16**	−.05
Number of users (4)					−.05	.09	.29***	.12*	.04	.29***	−.15**	.28***	.00
Average tweets per user (5)						.08	.07	−.11*	.04	−.06	−.16**	.03	−.12*
Household income (6)							.88***	−.44***	.43***	.14**	.03	−.16**	−.34***
% With bachelor's degree/higher (7)								−.46***	.47***	.13**	−.09	−.07	−.42***
Unemployment rate (8)									−.44***	.47***	−.10*	.31***	.39***
% With health-care coverage (9)										−.37***	.50***	−.31***	−.24***
Racial heterogeneity (10)											−.34***	.44***	.21***
Residence tenure (11)												−.19***	.08
Partisan polarity (12)													.03
Incivility (13)													

*$p < .05$, **$p < .01$, ***$p < .001$.

Hypothesis 2b suggested that residential tenure would be negatively related to incivility on Twitter. In the model consisting only of the control variables and social capital indicators, the residential tenure measure was related to incivility but in the opposite direction expected, $\beta = .16$, $p < .01$. This relationship remained significant after controlling for the effects of all other independent variables, $\beta = .13$, $p < .05$. Thus, Hypothesis 2b was not supported.

Finally, Hypothesis 3 suggested that districts with lower levels of partisan polarity would yield higher levels of uncivil Twitter activity. As seen in the polarity-only model, the relationship between partisanship and incivility was not significant. However, after controlling for the influence of the SES and social capital indicators, there was a negative and significant relationship between district partisanship and political incivility on Twitter, $\beta = -.18$, $p < .001$. This result was generally supportive of Hypothesis 3. Table 3 provides full description of the above-delineated results.[4]

Discussion

Using a big data approach, the present study drew upon a corpus of nearly 70 million tweets posted around the time of the 2012 presidential election. Given the scale of the data under consideration, this study developed and applied a computer-automated technique for the identification of online incivility. The present work also combined thousands of data points from the U.S. Census and voting outcomes from each U.S. district to explore the relationship between physical context and online behavior. In a broad sense, this study shows how social media data can be used in concert with traditional data to explain behaviors on a large scale.

Specifically, we investigated the relationship between district-level SES, social capital, and partisanship characteristics and the average amount of Twitter incivility that emanated from each U.S. congressional district. The currently observed results suggest that online incivility is positively associated with racial heterogeneity and negatively associated with district-wide education level and partisan polarity. When considered jointly, education level ($\beta = -.50$) emerged as the strongest predictor of district-wide incivility, followed by racial heterogeneity ($\beta = .28$) and partisan polarity ($\beta = -.18$).

There are a number of implications that stem from the current results. First, the strong relationship between district-wide education level and political incivility on Twitter is congruent with previous research, which has repeatedly identified education as an antecedent to political participation (e.g., Muhlberger, 2004; Verba, Scholzman, & Brady, 1995). Indeed, our results would seem to concur with the notion that "educational attainment is the primary mechanism behind many citizenship characteristics" (Hillygus, 2005, p. 25). Notably, prior research on political discussion has generally sought to associate education with involvement. The current results indicate that educational attainment may influence both the quantity/frequency of political discussion and the *quality* of such discussion.

Second, the results suggest that the social resources obtained off-line may be tied to the individual and social value gained from online engagement. While it may be the case that the Internet has the potential to fulfill social needs and enrich individual lives, it may also be the case that factors resultant in individual citizens' lived experiences may actively inhibit the fulfillment of such potential. While previous research has explored this from a knowledge acquisition standpoint (*the rich get richer hypothesis*; see Brundige & Rice, 2010, for review), our results suggest that the Internet may also play host to a *civility divide*. In other words, those equipped with economic and social privilege in the off-line realm may disproportionally gain value from online deliberation, while those with diminished economic and social resources may interact in a hostile, uncivil, and ultimately, less rewarding strata of the Internet.

Taken as a whole, the initial evidence presented here suggests that the tenor of online political discussion may be reflective of the broader social, cultural, and partisan contexts within which users physically reside (e.g., Dahlgren, 2005). Specifically, our analyses suggested factors that have been

Table 3. Standardized Coefficients Describing Relationship Between Social Capital Potential Indicators, SES Indicators, Partisan Polarity, and Twitter Incivility.

Variable Name	Controls Only	SES Indicators	Social Capital	Partisan Polarity	All Variables	VIF (All Variables Model)
District population	−.14**	−.07	−.12*	−.14**	−.07	1.13
Median age	−.25***	−.13*	−.20***	−.26***	−.13*	1.99
% Urban	−.11*	.05	−.21***	−.11*	−.01	2.34
Number of users	.00	.10*	−.04	.01	.10*	1.45
Average tweets per user	−.09	−.06	−.05	−.09	−.02	1.12
Household income		.25**			.08	5.89
% With bachelor's degree/higher		−.62***			−.50***	7.76
Unemployment rate		.20***			.11	2.43
% With health-care coverage		.10			.05	2.57
Racial heterogeneity			.32***		.28***	2.82
Residence tenure			.16**		.13*	2.14
Partisan polarity				−.04	−.18***	1.48
R^2	.08	.27	.15	.08	.32	
ΔR^2 (relative to controls only)		.19***	.07***	.001	.24***	
F	(5, 429) = 7.41 $p < .001$	(9, 425) = 17.45 $p < .001$	(7, 427) = 10.72 $p < .001$	(6, 428) = 6.28 $p < .001$	(12, 422) = 16.26 $p < .001$	

Note. VIF = variance inflation factor; SES = socioeconomic status.
*$p < .05$, **$p < .01$, ***$p < .001$.

previously associated with off-line incivility (i.e., racial heterogeneity and low levels of college education among the population) can similarly be correlated with online incivility. Understood in its broadest sense, our findings seem to offer support for theoretical perspectives that connect online behavior with lived experience. Moreover, in regards to the online potential of social/new media, these findings mirror others (e.g., Albrecht, 2006; Baek, Wojcieszak, & Delli Capini, 2011) in their suggestion that access cannot, alone, overcome the systematic inequalities that have traditionally structured political involvement.

Our results suggested that uncivil discourse was highest in districts that were characterized, in part, by factors traditionally thought to be indicative of a healthy and diverse democracy (i.e., low levels of partisan polarity and high levels of racial diversity). Clearly, these findings do not suggest that society is forced to choose between either a civil democracy or a well-functioning one. High levels of unconstrained and unfocused incivility are unlikely to yield the type of social cohesion and ideological compromise necessary for the continued and virile existence of the democratic state. That said, the present results suggest that there may be something of a misalignment between the idealized notion of democracy and its functional enactment.

Notably, we failed to either fully or partially support a number of our hypotheses. In some cases, such as those hypotheses related to SES, it could be that a single indicator (i.e., education level) served as the optimal representative of the on-hand phenomenon and, therefore, explained a bulk of the variance. In other cases, it could be that variable inter-relationships are highly conditional in nature. For instance, we found that residential tenure was positively—rather than negatively—related to incivility, suggesting that, in some cases, floating populations and inflows of new residents might actually serve as a dynamic force for civil discourse.[5] Future research should seek to explicitly and rigorously explore the degree to which social and economic factors exert moderating influences on the relationship between factors such as social capital, partisan polarity, and uncivil communication practices on Twitter.

A number of limitations temper the present findings. First, the nature of the data severely limits the generalizability of our findings. The source of data here, Twitter, is, at best, an instantaneous measure of behavior, not a durable measure of emotion or feelings (Vieweg, 2010). Moreover, Twitter cannot be reasonably understood to be a directly reliable proxy for public opinion in general. Also, the corpus here was limited to a specific event, the 2012 general election. The messages gathered in this analysis were also directed at a specific political candidate (e.g., Obama and/or Romney). While the findings still yield important conclusions toward discourse, democracy, and general elections, we cannot use the current results to make generalizations about the state of political discussion as a whole (either on or off of Twitter).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Notes

1. *Potential* could be interpreted in a number of ways. For instance, high levels of racial heterogeneity could present the potential for citizens to build ties with members of a number of different racial groups. However, in the present study, we interpreted potential in terms of factors previously shown to *empirically* affect the production and maintenance of social capital on the community/contextual level.
2. Reported correlations are from the disaggregated data set ($n = 143{,}404$).

3. Although researchers have previously outlined some advantages to use of geotagged Twitter data (e.g., Guo & Chen, 2014), questions relative to the degree that the geotagged subsample replicated the larger sample were addressed before proceeding to hypothesis testing. As the current sample was a subsample of a larger body of tweets on the 2012 election (over 70 million tweets from more than 11 million users), we were able to compare the characteristics of the geotagged subsample to the larger sample comprised of those users who opted out of geotagging. Examination of the raw individual-level data contained in each data set suggested that the distributional characteristics of the incivility variable were quite similar across samples (geotagged sample: $M = 0.25$, $SD = 0.52$, skewness = 6.01, kurtosis = 88.73; nongeotagged sample: $M = 0.24$, $SD = 0.66$, skewness = 7.46, kurtosis = 110.98). Reasoning that SES factors may be associated with the decision to opt-out of Geotagging (e.g., Hargittai, 2002; Peter & Valkenburg, 2006), we next performed a series of analyses to ensure that socioeconomic factors did not confound our results. Using the aggregated (district-level) data set, we calculated a variable that represented the percentage of geotag-enabled user accounts (total number of users per district in the current subsample) relative to the overall population. Percentages ranged from 0.01% to 0.20% ($M = 0.05\%$, $SD = 0.02\%$; skewness = 2.14, kurtosis = 8.44). We next correlated this variable with SES (percent of district with bachelor's degree or higher, unemployment rate, mean household income, and percent with health insurance) and control (median district age and urbanicity). Here we identified a positive correlation with education ($r = .27$, $p < .01$), a negative correlation with median district age ($r = -.14$, $p < .01$), and a positive correlation with urbanicity ($r = .23$, $p < .01$) factors. All other relationships were nonsignificant at $p < .05$. Finally, we estimated a series of nested ordinary least square regressions that employed incivility as the criterion. In the first model, we assessed the effects of control and SES factors. In the second model, we added the variable representing the percentage of geotag-enabled user accounts relative to the overall population. The results of these analyses suggested that addition of the latter variable had a negligible overall effect on the relationship between SES and civility, thus leading us to conclude while the decision not to opt out of geotagging may be connected to SES, this relationship should have a generally negligible effect on the hypotheses of interest in the current study.
4. Due to the strong correlation between the income and education measures ($r = .88$; see Table 2), the SES-only and all-variable models (Table 3) possessed relatively high levels of multicollinearity. Removal of the household income measure from these models did not meaningfully impact either the patterns of significance (including direct and relationship strength) or the amount of variance explained in the criterion variable. Moreover, the highest observed variance inflation factor coefficient was well below the commonly employed heuristic of 10.00 (e.g., O'Brien, 2007). As such, the models were presented in the initially hypothesized form.
5. Supplementary ordinary least square moderation analyses failed to indicate that SES factors exerted a moderating influence on the relationship between neighborhood tenure and uncivil discourse on Twitter. However, this finding may due to the relative lack of granularity in our unit of analysis. Future research should explore this relationship using neighborhood-level data.

References

Aiyer, S. M., Zimmerman, M. A., Morrel-Samuels, S., & Reischl, T. M. (2015). From broken windows to busy streets: A community empowerment perspective. *Health Education & Behavior*, *42*, 137–147.

Albrecht, S. (2006). Whose voice is heard in online deliberation? A study of participation and representation in political debates on the Internet. *Information, Communication & Society*, *9*, 62–82.

Alesina, A., & La Ferrara, E. (2002). Who trusts others? *Journal of public economics*, *85*, 207–234.

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The 'nasty effect': Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, *19*, 373–387.

Baek, Y. M., Wojcieszak, M., & Delli Capini, M. X. (2011). Online versus face-to-face deliberation: Who? Why? What? With what effects? *New Media & Society*, *14*, 363–383.

Baum, A., Garofalo, J. P., & Yali, A. M. (1999). Socioeconomic status and chronic stress: Does stress account for SES effects on health? *Annals of the New York Academy of Sciences*, *896*, 131–144.

Bimber, B., Cunill, M. C., Copeland, L., & Gibson, R. (2015). Digital media and political participation: The moderating role of political interest across acts and over time. *Social Science Computer Review*, *33*, 21–42.

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*, 117–136.

Borah, P. (2014). Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication Research*, *41*, 809–827.

Boyd, R. (2006). The value of civility? *Urban Studies*, *43*, 863–878.

Brady, D. W., & Han, H. C. (2006). Polarization then and now: A historical perspective. In P. S. Nivola & D. W. Brady (Eds.), *Red and blue nation* (Vol. 1, pp. 119–174). Baltimore, MD: The Brookings Institution.

Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review*, *89*, 271–294.

Brundige, J., & Rice, R. E. (2010). Political engagement online: Do the information rich get richer and the like-minded more similar? In A. Chadwick & P. N. Howard (Eds.), *Handbook of Internet politics* (pp. 144–156). New York, NY: Routledge.

Byrne, D. E. (1971). *The attraction paradigm*. New York, NY: Academic Press.

Campbell, C., Rockman, B. A., & Rudalevige, A. (2008). *The George Bush legacy*. Washington, DC: CQ Press.

Cappella, J. N., Price, V., & Nir, L. (2002). Argument repertoire as a reliable and valid measure of opinion quality: Electronic dialogue during campaign 2000. *Political Communication*, *19*, 73–93.

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

Cheng, H., & Riffe, D. (2008). Attention, perception, and perceived effects: Negative political advertising in a battleground state of the 2004 presidential election. *Mass Communication and Society*, *11*, 177–196.

Cho, J. (2013). Campaign tone, political affect, and communicative engagement. *Journal of Communication*, *63*, 1130–1152.

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, *64*, 658–679.

Coffey, D. J., Kohler, M., & Granger, D. M. (2015). Sparking debate campaigns, social media, and political incivility. In V. A. Farrar-Myers (Ed.), *Controlling the message: New media in American political campaigns* (pp. 245–269). New York: New York University Press.

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology*, *58*, 175–199.

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

Dahlberg, L. (2001). The Internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication & Society*, *4*, 615–633.

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

Dahlgren, P. (2003). Reconfiguring civic culture in the new media milieu. In J. Corner & D. Pels (Eds.), *Media and political style: Essays on representation and civic culture* (pp. 151–170). London, England: Sage.

Dahlgren, P. (2005). The Internet, public spheres, and political communication: Dispersion and deliberation. *Political Communication*, *22*, 147–162.

Davis, R. (2009). *Typing politics: The role of blogs in American politics*. New York, NY: Oxford University Press.

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have Americans' social attitudes become more polarized? *American Journal of Sociology*, *102*, 690–755.

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015, January 9). *Demographics of key social networking platforms*. Retrieved from Pew research Internet project http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/

Eveland, W. P., & Hively, M. H. (2009). Political discussion frequency, network size, and ''heterogeneity'' of discussion as predictors of political knowledge and participation. *Journal of Communication*, *59*, 205–224.

Freelon, D. (2013). Discourse architecture, ideology and democratic norms in online political discussion. *New Media & Society*, *17*, 772–791.

Guo, D., & Chen, C. (2014). Detecting non-personal and spam users on Geo-tagged Twitter Network. *Transactions in GIS*, *18*, 370–384.

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

Ellison, N. B., Steinfeld, C., & Lampe, C. (2006, June). *Spatially bounded online social networks and social capital: The role of Facebook*. Paper presented at the Annual Meeting of the International Communication Association, Dresden, Germany.

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

Farnam, T. W. (2012, February 12). Study: Negative campaign ads much more frequent, vicious than in primaries past. *The Washington Post*. Retrieved from http://www.washingtonpost.com/politics/study-negative-campaign-ads-much-more-frequent-vicious-than-in-primaries-past/2012/02/14/gIQAR7ifPR_story.html

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*, 563–588.

Forni, P. M. (2011, January 28). Why civility is necessary for society's survival. *The Dallas Morning News*. Retrieved from http://www.dallasnews.com/opinion/sunday-commentary/20100723-p.m.-forni-why-civility-is-necessary-for-society_s-survival.ece

Gil de Zúñiga, H., Jung, N., & Valenzuela, S. (2012). Social media use for news and individuals' social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication*, *17*, 319–336.

Gittell, R., & Vidal, A. (1998). *Community organizing: Building social capital as a development strategy*. Thousand Oaks, CA: Sage.

Glascock, J. (2014). Contribution of demographics, sociological factors, and media usage to verbal aggressiveness. *Journal of Media Psychology*, *26*, 92–102.

Goel, S., Anderson, A., Hofman, J., & Watts, D. (2015). The structural virality of online diffusion. *Management Science*, 1–17. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2015.2158

Goldstein, K., & Freedman, P. (2002). Campaign advertising and voter turnout: New evidence for a stimulation effect. *The Journal of Politics*, *64*, 721–740.

Goldstein, K., & Freedman, P. (2010). Lessons learned: Campaign advertising in the 2000 elections. *Political Communication*, *19*, 5–28.

Grand, S. R. (2014). *Understanding Tahir Square: What transitions can teach us about the prospects for Arab democracy*. Washington, DC: The Brookings Institution.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*, 1360–1380.

Gurin, P., Nagda, B. A., & Lopez, G. E. (2004). The benefits of diversity in education for democratic citizenship. *Journal of Social Issues*, *60*, 17–34.

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, 7. Retrieved from http://firstmonday.org/article/view/942/864

Heck, K. E., & Parker, J. D. (2002). Family structure, socioeconomic status, and access to healthcare for children. *Health Services Research*, *37*, 171–184.

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.

Hetherington, M. J. (2001, September). Resurgent mass partisanship: The role of elite polarization. In *American Political Science Association* (Vol. *95*, No. 03, pp. 619–631). Cambridge University Press: New York, NY.

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.

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

Hyman, H. (1959). *Political socialization*. Glencoe, IL: Free Press

Israel, G. D., Beaulieu, L. J., & Hartless, G. (2001). The influence of family and community social capital on educational achievement. *Rural Sociology*, *66*, 43–68.

Kan, K. (2007). Residential mobility and social capital. *Journal of Urban Economics*, *61*, 436–457.

Kang, N., & Kwak, N. (2003). A multilevel approach to civic participation: Individual lengths of residence, neighborhood residential stability, and their interactive effects with media use. *Communication Research*, *30*, 80–106.

Kasarda, J. D., & Janowitz, M. (1974). Community attachment in mass society. *American sociological review*, *39*, 328–339.

Klinkner, P. A., & Hapanowicz, A. (2005, July). Red and blue Déja Vu: Measuring political polarization in the 2004 election. In *The Forum* (Vol. *3*, No. 2).

Kavanaugh, A. L., Reese, D. D., Carroll, J. M., & Rosson, M. B. (2005). Weak ties in networked communities. *The Information Society*, *21*, 119–131.

Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. In N. Nohria & R. Eccles (Eds.), *Networks and organizations: Structure, form and action* (pp. 216–239). Boston, MA: Harvard Business School Express.

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.

Lau, R., Sigelman, L., Heldman, C., & Babbit, P. (1999). The effects of negative political advertisements: A meta-analytical assessment. *American Political Science Review*, *93*, 851–976.

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.

Liu, Q. A., Ryan, V., Aurbach, H., & Besser, T. (1998). The influence of local church participation on rural community attachment. *Rural Sociology*, *63*, 432–450.

Lutz, C., Hoffmann, C. P., & Meckel, M. (2014). Beyond just politics: A systematic literature review of online participation. *First Monday*, 19. Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/5260/4094

Marinov, N., & Schimmelfennig, F. (2015). *Does social media promote civic activism? Evidence from a field experiment*. Paper presented at the Spring Colloqium of the Center for Comparative and International Studies, Zürich, Germany.

McFarland, D., & Thomas, R. (2006). Bowling young: How youth voluntary associations influence adult political participation. *American Sociological Review*, *71*, 401–425.

McIntosh, H., Hart, D., & Youniss, J. (2007). The influence of family political discussion on youth civic development: Which parent qualities matter? *PS: Political Science and Politics*, *40*, 495–499.

McLeod, D. M., & Perse, E. M. (1994). Direct and indirect effects of socioeconomic status on public affairs knowledge. *Journalism and Mass Communication Quarterly*, *71*, 433–442.

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.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444.

Muhlberger, P. (2004). Access, skill, and motivation in online political discussion: Testing cyberrealism. In P. M. Shane (Ed.), *Democracy online: The prospects for political renewal through the Internet* (pp. 225–238). New York, NY: Routledge.

Mueller, A. (2012, September 21). *Detecting insults in social commentary*. Retrieved from https://www.kaggle.com/c/detecting-insults-in-socialcommentary/forums/t/2744/what-did-you-use/15951

Newton, K. (2001). Trust, social capital, civil society, and democracy. *International Political Science Review*, *22*, 201–214.

Nunn, L. M., & Evans, J. H. (2006). Geographic polarization in social attitudes. In Annual meeting of the American Sociological Association, Montreal Convention Center. Quebec, Canada.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*, 673–690.

Oliver, J. E. (1999). The effects of metropolitan segregation on local civic participation. *American Journal of Political Science*, *43*, 186–212.

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

Oropesa, R. S. (1992). Social structure, social solidarity and involvement in neighborhood improvement associations. *Social Inquiry*, *62*, 107–118.

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

Peter, J., & Valkenburg, P. (2006). Adolescents Internet use: Testing the 'disappearing digital divide' versus the 'emerging digital differentiation' approach. *Poetics*, *4-5*, 293–405.

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

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

Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society*, *18*, 121–138.

Rubenstein, L. (2014, September 16). 2014 general election advertising opens even more negative than 2010 or 2012. *Wesleyan Media Project*. Retrieved from http://mediaproject.wesleyan.edu/releases/2014-general-election-advertising-opens-even-more-negative-than-2010-or-2012/

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

Sadilek, A., Kautz, H. A., & Silenzio, V. (2012). Predicting disease transmission from Geo-Tagged Micro-Blog data. *Paper presented at the Twenty-Sixth AAAI Conference on Artificial Intelligence* (pp. 136–142). Retrieved from http://www.aaai.org/Conferences/AAAI/aaai12.php

Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice*, *8*, 18–33.

Settle, J. E., Bond, R. M., Coviello, L., Fariss, C. J., Fowler, J. H., & Jones, J. J. (2014). From posting to voting: The effects of political competition on online political engagement. *Political Science Research and Methods*, *2004*, 1–18.

Settle, J. E., Bond, R., & Levitt, J. (2011). The social origins of adult political behavior. *American Politics Research*, *39*, 239–263.

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., 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.

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. Chicago, IL: University of Chicago Press.

Steptoe, A., & Feldman, P. J. (2001). Neighborhood problems as sources of chronic stress: Development of a measure of neighborhood problems, and associations with socioeconomic status and health. *Annals of Behavioral Medicine*, *23*, 177–185.

Stromer-Galley, J., & Muhlberger, P. (2009). Agreement and disagreement in group deliberation and its consequences: Satisfaction, engagement, and opinion reevaluation. *Political Communication*, *26*, 173–192.

Sunlight Congress API. (2014, January 12). *Official API Documentation*. Retrieved from https://sunlightlabs.github.io/congress

Sunstein, C. (2001). *Republic.com*. Princeton, NJ: Princeton University Press

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, *61*, 2544–2558.

Turner, J. C. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, England: Basil Blackwell.

United States Census Bureau. (2014, December 4). *2013 ACS Data Release*. Retrieved from http://www.census.gov/acs/www/data_documentation/2013_release/

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 students' life satisfaction, trust, and participation. *Journal of Computer-Mediated Communication*, *14*, 875–901.

Vargo, C. (2014). *Brand messages on Twitter: Predicting Diffusion with Textual Characteristics* (Doctoral dissertation). The University of North Carolina at Chapel Hill.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

Vieweg, S. (2010). *The ethics of Twitter research*. Paper presented at the Computer Supported Cooperative Work 27. (pp. 1–3), Savannah, GA.

Vissers, S., & Stolle, D. (2014). The Internet and new modes of political participation: Online versus offline participation. *Information, Communication & Society*, *17*, 937–955.

Wasserman, D. (2013). *Introducing the 2014 Cook Political Report Partisan Voter Index*. Retrieved December 12, 2014, from http://cookpolitical.com/story/5604

Wegner, G. (1991). Review: Political socialization and schooling. *History of Education Quarterly*, *31*, 389–391.

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.

Wolak, J. (2006). The consequences of presidential battleground strategies for citizen engagement. *Political Research Quarterly*, *59*, 353–361.

Yuan, Y. C., & Gay, G. (2006). Homophily of network ties and bonding and bridging of social capital in computer-mediated distributed teams. *Journal of Computer-Mediated Communication*, *11*, 1062–1084.

Zamith, R., & Lewis, S. C. (2015). Content analysis and the algorithmic coder what computational social science means for traditional modes of media analysis. *The ANNALS of the American Academy of Political and Social Science*, *659*, 307–318.

Author Biographies

Chris J. Vargo, PhD, is an assistant professor at the University of Alabama in the College of Communication and Information Sciences. He applies computer science methodology to social science theory. He can be reached at cjvargo@ua.edu

Toby Hopp, PhD, is an assistant professor at the University of Alabama in the College of Communication and Information Sciences. He researches civility as it pertains to democratic activities. He can be reached at hopp@apr.ua.edu