Political discourse on social networking sites: Sentiment, in-group/out-group orientation and rationality

Scott P. Robertson*, Sara Douglas, Misa Maruyama and Bryan Semaan

Information and Computer Sciences Department, University of Hawaii, Honolulu, HI, USA

Abstract. The news feeds of two U.S. politicians’ Facebook sites were examined across 22 months leading up to an election in order to explore changes in social-network-mediated public political discourse over time. Changes over time were observed in who was being addressed and in the affective valence of comments. A complex flow of attention between in-group and out-group concerns was observed with in-group comments dominant both in early and late phases. Also, positive comments decreased and negative comments increased over time. These phenomena, dubbed “reflection-to-selection” and “converging sentiment”, were refined to explain the observed nonlinearities. The flow of rational versus affective comments in politicians’ Facebook data across time was also explored. Comments reflecting cognition were more prevalent at all times than comments reflecting affect, but their distribution also varied in complex ways over time. Finally, the concept of “potential public sphere” in contrast to “realized public sphere” in virtual spaces is introduced.

Keywords: Digital government, social networking, e-participation, e-citizenship, virtual public sphere

1. Introduction

1.1. The internet and political activity

Since social networking sites (SNSs) first appeared on the Internet, their use has consistently increased and the purposes for which they are used have broadened. A recent survey conducted by the Pew Internet and American Life project [37] reports that 65% of all Americans say they use SNSs, up from 2% since 2005. A 2011 Nielsen study [47] reports that Americans spend 23% of their Internet time using blogs and social networks, a larger percentage than any other online activity. Americans spend more time on Facebook than any other site, and Facebook is the most popular application across all ICT platforms [21].

An increasingly important use of SNSs is for participation in civic and political activities. Hampton et al. [21] report that, controlling for demographic characteristics such as education and income, Internet users are much more likely than non-Internet users to have been involved in civic/political activities such as attending a political rally (2.39 times more likely), trying to influence a vote (75% more likely), and...
2. S.P. Robertson et al. / Political discourse on social networking sites

Actually voting or reporting intention to vote (53% more likely). There has been considerable investigation and controversy about whether Internet use generally, and SNS use specifically, has an impact on civic participation [1,12,48,49], but data shows that SNS use and civic activity are certainly related since frequent Facebook users, when compared to other Internet users, are 2.5 times more likely to attend rallies, 57% more likely to have tried to influence a vote, and 43% more likely to have reported voting or intending to vote [21].

1.2. Goals of this article

In this article we explore comments posted on the Facebook sites of politicians running for office. This is an area where civic discourse should be prevalent. Our goal is to describe the types of discourse that occur in the virtual public sphere of SNSs and show how this discourse changes over time. We re-examine two previous phenomena that we have observed and reported elsewhere [59,63], reflection-to-selection and converging sentiment, to find more subtle changes over time. Further, we examine rational and affective components of the dialogue taking place on politicians’ official Facebook pages. We ask whether both take place, in what proportions, and whether there are changes over time.

2. Literature review

2.1. ICSs and the public sphere

The social networking components of many online sites are often described as a new type of socio-technical public sphere such as Habermas [7,20] described in non-Internet contexts. Researchers have asked whether true civic discourse takes place online [12,14,16,17] and whether the notion of a public sphere is changed by socio-technical mediation [2,4,6,8,9,75]. Recent empirical work has sought to document as public spheres many types of online spaces such as political blogs [10,15,70,73] and microblogs [41], discussion forums [52], newsgroups [24], online videos [44,45], and SNSs [32,40,42,61,62,78,79]. The ideal of a public sphere for political discourse is that it will support rational discussion in which multiple viewpoints are heard and consensus is reached, although such ideal discussions may rarely occur in real or virtual contexts.

2.2. Why people use SNSs

When we examine discourse over time, we must ask what brings people back or what sustains their conversation. Joinson [26] provides an early study of the reasons that people give for using Facebook. He found seven prominent motivations:

– Keeping in touch,
– Interacting with others who have shared interests and identities,
– Viewing photographs,
– Using applications like games,
– Finding new friends and acquaintances through social investigation,
– Browsing friends of friends, called “social network surfing”,
– Informing others via status updating.
The second motivation is related to joining groups, organizing activities, and meeting like-minded people and tends to be associated with greater time spent in SNSs. Lampe et al. [32] referred to this activity as “social browsing”. Interestingly, learning or discovering something outside of the arena of information about friends is not included among these motivations, except to the degree that it is found in the shared interests and identities motivation.

Lampe et al. [33] showed that the motivations of users changed during their involvement with groups. They emphasized that the motivations for joining a group differ from the motivations for remaining in a group. Often the change is from information seeking in the beginning to issues such as maintaining connectivity and gaining satisfaction and a sense of belonging in later stages. Dholokia et al. [13] distinguish among several motivations for participating in online communities: purposive value, self discovery, maintaining interpersonal connectivity, gaining social enhancement, and deriving entertainment. These vary across different types of communities and through time.

Ren et al. [57] distinguish between identity-based reasons for participating on online communities and bond-based reasons [53,54,71,72]. Common identity underlies users’ participation in communities when they feel that they have shared interests. Members of an identity-based group will not necessarily expand their relationships with others beyond the shared concern of the group. Common bond underlies users’ participation in communities when they are already bonded in some way such as friendship or family. Members of a bond-based group usually share common interests, but their participation in the group is not limited to the shared interest(s). Ren et al. [57] showed that the design of the socio-technical environment can differentially influence its usefulness for identity-based versus bond-based communities.

The degree to which social browsing operates to expose SNS users to new ideas will depend on the context in which the browsing takes place. Whether or not political discourse is broadened by exposure to diverse ideas in computer-mediated settings has been a subject of debate [5,18,25,67–69,74]. On the one hand, online discourse may broaden exposure to a diverse set of ideas [55,76,77,79]. On the other hand, it has also been observed to narrow exposure to ideas in an echo-chamber of like-minded friends (i.e. “homophily”) [1,32,43,46,48,49].

2.3. Rationality and emotion in political discourse

Political decision making has both rational and emotional components. A rational-actor model of political decision making emphasizes the gathering of information, weighing of choices, and reasoning about costs and benefits of various decisions. While such models have been prevalent in political science for many decades [58], they have been challenged by theorists and researchers coming from psychology and other behavioral social sciences [34,65,66].

One challenge involves the pragmatics of making complex decisions that require significant cognitive resources. Lau [34] points out that many people are unmotivated to put in the cognitive effort required to make complex decisions that can be made effectively enough by using short-cut heuristics. This view falls into a broader category of bounded rationality theories [23,27,64–66] in which actors weigh the complexity of a problem against their cognitive and other resources. A popular theory in political psychology is the “on-line running tally” idea that people create an initial, often schema-based concept of a candidate which they then use as a comparator when new information is encountered [35,36,56]. The candidate schema is updated as necessary while the information used for updating might be forgotten.

Social psychologists have spent considerable effort studying the role of emotion in the political arena [19,38]. Marcus et al. [39] propose that political judgment is a balance between habit and reason that is mediated largely by emotion. Studies of the impact of media on decision making have also
emphasized the role of impression formation and manipulation in forming political judgments [80]. In this paper, we begin an examination of rational and affective components of discourse in SNSs and ask how these two types of discourse change over time.

2.4. “Reflection-to-selection” and “converging sentiment”

Robertson and his colleagues [59–63,73] have examined whether the comment areas of politicians’ Facebook pages serve as socio-technical public spheres that encourage exploration of diverse ideas and mutual exchange of information [10,11]. A politicians’ comment area on a SNS supports an identity-based group. However, participants in the comment area are both in-group members (supporters of the candidate) and out-group members (opponents of the candidate). The overarching shared interest is an upcoming election. Within that broad interest group are opposing subordinate interest groups for each candidate.

In previous work [59,63], Robertson and his colleagues found evidence for several general types of changes in discourse on politicians’ Facebook sites over the months leading up to an election. These changes were related to group identity and to the valence of conversation. In virtually all cases, the use of first-person pronouns (“I” and “we”) and second person pronouns (“you”) decreased over time while the use of third-person pronouns (“he”, “she”, and “they”) increased over time. Table 1 shows the observed correlations of number of pronouns in these categories with time from the two studies. Further examination of the pronoun “you” showed that the decrease over time was for instances where the candidate was being directly addressed and not when “you” referred to other people. This pattern was interpreted as evidence for a “Reflection-to-Selection” model in which the community was inward facing and engaged in defining a shared identity during the early stages of a political campaign (the “Reflection” stage), but then later turned outward toward others and the opponent (the “Selection” stage).

A sentiment analysis of the affective nature of the comments showed that dialog tended to be more positive than negative, but that positive comments decreased over time while negative comments increased over time such that they were almost equal in proportion when the elections occurred. Table 2 shows the observed correlations of positive and negative affect terms with time from two studies. This pattern was interpreted as evidence for a “Converging Sentiment” phenomenon in which early identity definition was largely a positive activity focused on shared values and ideas but, once formed on both sides, turned into an identity contrast activity that was more negative.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>2008 U.S. presidential election</th>
<th>2010 California gubernatorial election</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obama</td>
<td>Clinton</td>
</tr>
<tr>
<td>First-person (I, we)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.87*</td>
<td>−0.52*</td>
</tr>
<tr>
<td>Third-person (s/he, they)</td>
<td>+0.84*</td>
<td>+0.80*</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>2008 U.S. presidential election</th>
<th>2010 California gubernatorial election</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obama</td>
<td>Clinton</td>
</tr>
<tr>
<td>Positive comments</td>
<td>−0.86*</td>
<td>−0.64*</td>
</tr>
<tr>
<td>Negative comments</td>
<td>+0.90*</td>
<td>+0.52*</td>
</tr>
<tr>
<td>Positive – negative</td>
<td>−0.90*</td>
<td>−0.64*</td>
</tr>
</tbody>
</table>
In this article we re-examine the data collected from the 2010 gubernatorial election in California [63]. Specifically, we looked for non-linear aspects of the pronoun and sentiment data. We also added observations about rational and non-rational discourse as reflected by words describing cognitive mechanisms and affect.

3. Method and data set

User comments from the Facebook news feeds of the two candidates in the 2010 California gubernatorial race, Jerry Brown and Meg Whitman, were examined for this study. Neither candidate was an incumbent although both were well known in California. Brown was serving as the State Attorney General of California at the time of the election and had been the Governor from 1975–1983. Whitman was a former CEO of eBay.

Public comments open to any Facebook user or authorized application were collected from the February 9, 2009 to November 13, 2010 time period. The election itself took place on November 2, 2010. Comments were extracted using Facebook’s Open Graph API in compliance with Facebook rules and Open Graph restrictions. There were 9,442 comments from 6,804 individuals collected from Jerry Brown’s Facebook site and 11,787 comments from 5,039 individuals collected from Meg Whitman’s Facebook site. Figure 1 shows the number of words appearing per month on the two candidates’ Facebook sites in the months leading up to the election.

Linguistic Inquiry and Word Count (LIWC) [50,51] was used to analyze the corpus. LIWC provides basic word counts and percentages of words in several part-of-speech and semantic categories. Semantic categories were derived empirically by the LIWC developers, and word categories have been psychometrically tested [51] and used previously in a variety of analyses of discourse [3,22,28–31].

The reflection-to-selection phenomenon was examined by using the person of pronouns (first-person, second-person, and third-person). The converging sentiment phenomenon was examined by finding the valence (positive or negative) of words reflecting affect. The rationality question was addressed by examining words in the cognitive mechanism category and words in the general affect category. Cognitive mechanism words are terms that reflect discussion of mental processes, insight, and analysis (e.g.,
“think”, “know”, “cause”). Affect words are terms that express emotion and feeling (e.g. “feel”, “like”, “hate”).

The ordinal sequence of months leading up to the election was coded with the numbers 1–22 where 1 is the farthest month preceding the election and 22 is the month of the election. Multiple regression analyses were performed using the linear time sequence as one predictor and its quadratic transformation as a second predictor. The quadratic trend in regression tests for a line which curves in one direction. Separate analyses were performed on the following criterion variables:

- Whitman’s use of first-person pronouns.
- Brown’s use of first-person pronouns.
- Whitman’s use of third-person pronouns.
- Brown’s use of third-person pronouns.
- Whitman’s use of words reflecting cognitive mechanism.
- Brown’s use of words reflecting cognitive mechanism.
- Whitman’s use of words reflecting affect.
- Brown’s use of words reflecting affect.
- Whitman’s use of words reflecting positive affect.
- Brown’s use of words reflecting positive affect.
- Whitman’s use of words reflecting negative affect.
- Brown’s use of words reflecting negative affect.

For the content analysis, a random sample of 100 posts was taken for each candidate from four different time periods. The time periods were determined by a combination of number of posts within the period (in order to guarantee a large enough sample pool in each period since the number of posts greatly increased in the months just before the election) and important events. The sample periods were February 9, 2009–December 31, 2009 (691 posts for Brown and 465 posts for Whitman covering the beginning of the Facebook activity through the end of the year), January 1, 2010–June 8, 2010 (2122 posts for Brown and 1901 posts for Whitman leading up to the primary election), June 9, 2010–September 28, 2010 (2667 posts for Brown and 3714 posts for Whitman), and September 29, 2010–November 13, 2010 (3971 posts for Brown and 5707 posts for Whitman immediately before and after the election). This resulted in a sample of 800 posts, 400 for each candidate, which were then categorized independently by two of the authors (SD and MM). Agreement between the experimenters was above 80% for each category. Disagreements were reconciled for all sampled posts.

4. Results

4.1. Reflection-to-selection

Table 3 shows the mean number of first-person and third-person pronouns used by both candidates combined across the 22 time periods in this study. A 2 × 2 repeated-measures ANOVA was performed using time period as if it were a subject factor. There was significantly more use of first-person pronouns (mean = 4.11 percent) than third-person pronouns (mean = 1.73 percent), $F(1,21) = 121.02, p < 0.0001$, and significantly more pronoun use by Whitman in contrast to Brown, $F(1,21) = 4.84, p < 0.039$. However, there was also an interaction between the pronoun variable and the candidate $F(1,21) = 5.48, p < 0.029$. Examination of the means suggests that although the percentage of third-person pronouns was about the same for both candidates (means = 1.68 and 1.78 percent for Whitman and
Table 3

<table>
<thead>
<tr>
<th>Pronoun person category</th>
<th>Meg Whitman</th>
<th>Jerry Brown</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-person</td>
<td>4.40</td>
<td>3.83</td>
<td>4.11</td>
</tr>
<tr>
<td>Third-person</td>
<td>1.70</td>
<td>1.78</td>
<td>1.73</td>
</tr>
<tr>
<td>Mean</td>
<td>3.04</td>
<td>2.80</td>
<td></td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Cognition/affect category</th>
<th>Meg Whitman</th>
<th>Jerry Brown</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive mechanism</td>
<td>15.95</td>
<td>14.21</td>
<td>15.08</td>
</tr>
<tr>
<td>Affect</td>
<td>5.66</td>
<td>5.66</td>
<td>5.66</td>
</tr>
<tr>
<td>Mean</td>
<td>10.80</td>
<td>9.93</td>
<td></td>
</tr>
</tbody>
</table>

Brown respectively), the percentage of first-person pronouns was higher for Whitman (mean = 4.40 percent) than for Brown (mean = 3.83 percent).

For Whitman the percentage of first-person pronouns decreased significantly with the linear time predictor, $b = -0.053, t(20) = -2.10, p < 0.049$, and this predictor explained a significant proportion of the variance, $R^2 = 0.18, F(1,20) = 4.42, p < 0.049$. Addition of the quadratic predictor did not increase the variance explained, $b = 0, t(19) < 1.0, ns$. For Brown the percentage of first-person pronouns also decreased significantly with the linear time predictor, $b = -0.064, t(20) = -3.00, p < 0.007$, and this predictor explained a significant proportion of the variance, $R^2 = 0.31, F(1,20) = 8.98, p < 0.007$.

Addition of the quadratic predictor marginally improved the prediction, $b = 0.007, t(19) = 2.02, p < 0.06$, and significantly increased the variance explained, $R^2 = 0.43, F(2,19) = 7.22, p < 0.005$.

For Whitman the percentage of third-person pronouns remained the same throughout the study period. For Brown, on the other hand, the percentage of third-person pronouns increased significantly with the linear time predictor, $b = 0.039, t(20) = 2.04, p < 0.05$, and this predictor explained a significant proportion of the variance, $R^2 = 0.17, F(1,20) = 4.17, p < 0.054$. Addition of the quadratic predictor did not reach significance, $b = -0.006, t(19) = -1.79, p < 0.09$, but with the two predictors the proportion of variance explained was considerably increased, $R^2 = 0.29, F(2,19) = 3.91, p < 0.038$.

4.2. Cognition and affect

Table 4 shows the mean number of words in the cognition (the “cognitive mechanism” category in LIWC) and affect categories for both candidates combined across the 22 time periods in this study. A 2 × 2 repeated-measures ANOVA was performed using time period as if it were a subject factor. There was significantly more use of terms related to cognition (mean = 15.08 percent) than affect (mean = 5.66 percent), $F(1,21) = 2126.09, p < 0.0001$, and significantly more terms of both types used by Whitman in contrast to Brown, $F(1,21) = 19.30, p < 0.0001$. However, there was also an interaction between the cognition/affect variable and the candidate $F(1,21) = 24.44, p < 0.0001$. Examination of the means suggests that although the percentage of affect terms was the same for both candidates (mean = 5.66 percent for Brown and also for Whitman), the percentage of cognition terms was slightly higher for Whitman (mean = 15.94 percent) than for Brown (mean = 14.21 percent).

Figure 3 shows the percentage of comments that were categorized by LIWC as being related to cognition and the number of comments that were categorized as being related to affect over the 22 months leading to the election. The best-fitting linear and quadratic lines are drawn on the plots.

For Whitman the percentage of words related to cognitive mechanism decreased marginally with the linear time predictor, $b = -0.051, t(20) = -1.99, p < 0.06$, and this predictor explained a marginally
significant proportion of the variance, \( R^2 = 0.17, F(1,20) = 3.97, p < 0.06 \). Addition of the quadratic term produced a much better prediction, \( b = -0.011, t(19) = -2.99, p < 0.008 \), and considerably increased the variance explained, \( R^2 = 0.43, F(2,19) = 7.24, p < 0.005 \). For Brown the percentage of words related to cognitive mechanism remained the same throughout the study period.

For Whitman the percentage of words related to affect was not predicted by the linear time predictor, \( b = -0.024, t(20) = -0.84, ns \), but addition of the quadratic term produced a significant prediction, \( b = 0.010, t(19) = 2.23, p < 0.038 \). Despite this, the proportion of variance explained by the two terms was not significant, \( R^2 = 0.23, F(2,19) = 2.90, p < 0.079 \). For Brown the percentage of words related to affect remained the same throughout the study period.
Table 5
Mean percentage of positive and negative sentiment words in two political candidates’ Facebook news feeds combined across 22 months preceding an election

<table>
<thead>
<tr>
<th>Sentiment category</th>
<th>2010 California gubernatorial candidates</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Meg Whitman</td>
<td>Jerry Brown</td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>Positive sentiment</td>
<td>4.29</td>
<td>3.92</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>Negative sentiment</td>
<td>1.45</td>
<td>1.67</td>
<td>1.56</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.87</td>
<td>2.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. Linear and quadratic trends for the percentage of “cognitive mechanism” and “affect” terms in the news feeds of Meg Whitman and Jerry Brown in the months leading up to the 2010 California Gubernatorial election.
4.3. Converging sentiment

Table 5 shows the mean number of words in the positive and negative affect categories of LIWC for both candidates combined across the 22 time periods in this study. A $2 \times 2$ repeated-measures ANOVA was performed using time period as if it were a subject factor. There was significantly more use of words reflecting positive affect (mean = 4.10 percent) than words reflecting negative affect (mean = 1.56 percent), $F(1,21) = 296.88$, $p < 0.0001$. There were no other significant effects.

Figure 4 shows scatterplots of the percentage of positive and negative sentiment terms in the comments...
For Whitman the percentage of words reflecting positive sentiment decreased significantly with the linear time predictor, $b = 0.067$, $t(20) = -2.43$, $p < 0.025$, and this predictor explained a significant proportion of the variance, $R^2 = 0.23$, $F(1,20) = 5.90$, $p < 0.05$. Addition of the quadratic predictor did not increase the variance explained, $b = -0.005$, $t(19) = 1.11$, ns. For Brown the percentage of words reflecting positive sentiment was not significantly related to time using the linear predictor alone, however addition of the quadratic component resulted in a significant prediction, $b = -0.012$, $t(19) = -2.48$, $p < 0.011$ and explained a significant proportion of variance, $R^2 = 0.32$, $F(2,19) = 4.47$, $p < 0.05$.

For Whitman the percentage of words reflecting negative sentiment increased significantly with the linear time predictor, $b = 0.043$, $t(20) = 3.44$, $p < 0.003$, and this predictor explained a significant proportion of the variance, $R^2 = 0.37$, $F(1,20) = 11.86$, $p < 0.003$. Addition of the quadratic predictor did not increase the variance explained, $b = 0.004$, $t(19) = 1.77$, ns. For Brown the percentage of words reflecting negative sentiment was not significantly related to time using the linear predictor alone, however addition of the quadratic component resulted in a significant prediction, $b = -0.006$, $t(19) = -2.40$, $p < 0.027$ and explained a (weakly) significant proportion of variance, $R^2 = 0.326$, $F(2,19) = 3.26$, $p < 0.061$.

### 4.4. Addressee and content

Based on prior studies [60,62] and informal sorting data from subjects, the following six comment categories were coded:

- Direct address to the candidate.
  - “I’ll support you and I hope you support my 2nd amendment rights and be tough on immigration.”
  - “America Needs More SOUL. I believe in you. Go Jerry Go!!”

- Direct address to another commenter
  - “I think this lady has what it takes. Look at her history, she has been a success everywhere she’s been.”
  - “This woman has nothing to say. Its all meaningless slogans.”

- Evidence
  - “Meg Whitman Just turned coat on immigration. She just put an ad on Spanish TV networks saying she was against Arizona’s 1070, and Proposition 187, A law that would have denied Illegal immigrants access to public benefits. TWO FACED!”
  - “The New Jim Crow www.huffingtonpost.com There are more African Americans under correctional control today – in prison or jail, on probation or parole – than were enslaved in 1850, a decade before the Civil War began.”

- Action
  - “Should Illegal Aliens be Granted In-State (Resident) Tuition Rates at State Colleges and Universities? Please vote and leave your comments: http://www.stopillegalaliens.info/”
  - “If you dont like Meg join this! http://www.facebook.com/home.php#!/pages/I-bet-this-pencil-can-get-more-fans-than-Meg-Whitman/110010405703205”

- Rebuttal
Fig. 5. Number of comments in each comment category across four time periods from Jerry Brown’s Facebook wall.

Fig. 6. Number of comments in each comment category across four time periods from Meg Whitman’s Facebook wall.

∗ “If the worse thing she’s done is not vote, shown the same apathy towards government that many of us has felt, while building a hugely successful business, I can live with that!”

∗ “This is getting old, there are more important things that should be on the court dockets then this! it’s time to let people live their lives the way they want as long as they don’t break any laws it’s not up to us to make this choice for them.”

– Joking and Ridicule

∗ “I have a fabulous idea for the Nutmeg political action team. Just find a way to hook this page to the Hyperion Plant in LA (wastewater treatment plant). Presto! A ready stream of already steaming statements!”

∗ “I just found out that my 4th grade daughter has been representing the ‘Jerry Brown voters’ in her class discussions about the governors race. I just hope she hasn’t been repeating everything she’s heard me say about Meg Whitman, or I’ll be getting a note from her teacher soon.”

The first two categories are particularly important with regard to the hypothesis that converging sentiment and reflection-to-selection involve a turn from inward-directed, community-building activity to
Fig. 7. Number of comments directed toward the candidate, the opponent, and others across four time periods from Jerry Brown’s Facebook wall.

Fig. 8. Number of comments directed toward the candidate, the opponent, and others across four time periods from Meg Whitman’s Facebook wall.

more outward-directed activity. Comments in the first category are relevant to identity-based participation [57] since they involve interaction with the object around which communality is being constructed. Comments in the second category can reflect community-building if they are positive and directed towards others within the community or negative and directed to outsiders. They are disruptive to the community if they are negative and directed to community members or positive and directed to outsiders.

Figures 5 and 6 show the number of comments in each category across the four time periods (out of 100 per time period) for the two candidates. The number of comments differed considerably across the categories, $\chi^2(5) = 831.58, p < 0.01$ and $\chi^2(5) = 719.37, p < 0.01$ for Brown and Whitman respectively, with direct address to candidates and direct address to other commenters dominating. In Brown’s case, direct address to the candidate decreased across time periods, $\chi^2(5) = 8.85, p < 0.05$.

Figures 7 and 8 show comments combined across all of the previous categories and resorted in terms of whether they were directed toward the candidate, the opponent, or others. Again, for Brown com-
Fig. 9. Number of comments in the five intention categories on the candidates’ Facebook walls.

Comments directed at the candidate decreased over time, \( x^2(3) = 7.86, p < 0.05 \), while comments directed at others increased over time, \( x^2(3) = 13.40, p < 0.05 \). Although there was a similar trend for Whitman, differences across time periods were not significant. There were very few comments within either candidate’s site that were directed at the opponent.

Direct address to the candidate constituted a large percentage of all comments. Users had many goals when they made a direct address. Categorization of comments within the direct address category was carried out focusing on the intention of the commenters. The following seven types of comments were identified:

- **Question**: Asking the candidate for his or her position on an issue.
- **Suggestion**: Telling the candidate that they should do or try something.
- **Request**: Asking the candidate to do something.
- **Support**: Expressions of enthusiasm and encouragement.
- **Invitation**: Asking the candidate to participate in something.
- **Stories**: Telling personal narratives or describing things that have happened to the poster.

Figure 9 shows the number of comments in each intention subcategory for both candidates. Trends over time varied dramatically but non-systematically across time periods, so the data was collapsed over time.

The distribution of comment types was unequal across categories and between candidates, \( x^2(7) = 36.7, p < 0.05 \). The most significant observations of differences between candidates were that Whitman was asked three times as many questions as Brown, but Brown received 7 times as many requests as Whitman. Brown also received considerably more support activity. Across categories, it is clear that one of the primary purposes of dialog on candidate’s websites is to offer support.

5. Discussion

In several instances over two elections we observed the *reflection-to-selection* phenomenon, which we initially described as a linear model in which commenters are focused inward when they begin to build their online community but turn outward with time. Our new analyses show that another pattern exists in which commenters return to an inward orientation at election time. This pattern was present...
In the data for the winner of the California Gubernatorial race (Brown’s data in Fig. 2), but not in the
loser’s data. It remains to be seen if this is replicable and, if so, explainable. It is interesting that the
inward-facing first-person pronoun and the outward-facing third-person pronoun seem to trade off. The
U-shaped function for first person pronoun data for Brown is mirrored by an inverted U-shape in the
third-person pronoun data. These two pronoun types are theoretically independent of each other, they
could go up or down together or be completely unrelated. However this effect suggests that commenters
are focused predominantly one way or the other at any given time and may drift from one focus to the
other as their goals change through time.

In several instances over two elections we observed the *converging sentiment* phenomenon, which we
also initially described as a linear model in which commenters are most positive and least negative when
they are first forming their community, but become less positive and more negative over time. Our new
analysis again shows that another pattern exists in which commenters may become more positive and
less negative once again at the end of their quest as the event that they have organized around draws
near. At least this pattern was present for the winner of the election. The community formed around
the loser of the election showed a different pattern in which they plateaued at a lower level of positive
commenting as the election neared, but accelerated the increase in comments reflecting negative affect.

In data first presented here, we found that political discourse in SNSs consists of both rational discus-
sion (as reflected by the *cognitive mechanism* category of LIWC) and affective commenting (as reflected
in the *affect* category of LIWC). Rational discussion is more prevalent than affective commenting, and
this trend remains the same over at least several months prior to the election. Whitman’s data showed a
tantalizing trend in which at first rational commenting increased through time and affecting commenting
decreased through time, but as the election neared this trend reversed such that rational commenting
began to decline and affective commenting began to increase. Like pronoun direction and affective va-
dence, these two types of discourse are independent and do not have to track each other reciprocally. The
observation that they do suggests that communities cannot simultaneously maintain both rational and
affective interaction. Perhaps a type of affect contagion, or rationality contagion, takes place to move the
whole community into one discourse context or the other. The direction of movement of a community
along these scales may say a lot about the health of the community at any given moment in time.

Content analysis of the comments showed a clear trend such that engagement with a community’s
candidate is dominant early in campaigns and engagement with others is dominant in later stages. One
explanation for this trend is that political communities built around candidates need to construct and
solidify the components of their common interest. This means initially asking many questions of the
candidate. This period tends to be highly positive in terms of comment sentiment. After the group is
more well defined, however, the communal interest becomes focused on the other candidate and the
competitive group(s). Community members evolve from a reflective mode into a more active mode that
is sensitive to external events and the ongoing contest. There becomes more of an emphasis on defining
the other, which in the context of politics generates more negative comments.

The predominance of cognitive words in the discourse on candidates’ Facebook pages is good news for
proponents of the view that SNSs can serve as a public sphere. Since one of the defining characteristics
of a public sphere is that it allows rational arguments to be aired and discussed openly, the presence of
these terms suggests that this activity is indeed taking place with high frequency. Of course, the presence
of affective terms also shows that public discourse on SNSs is passionate. While these types of discourse
may fluctuate in relative frequency, they are co-present at all times. Readers browsing SNS material must
skip through many types of comments and posts in order to experience the content. In this way SNSs are
unlike face-to-face public spheres in which attention must be managed more carefully and participants
are not free to navigate the information threads of their own interests.
SNS’s might better be described as user-generated “potential public spheres”, information-opportunity contexts that provide options for browsing, navigation and participation. They become “realized public spheres” only when people take action, for example by successfully finding interesting information, making reasoned decisions, contributing comments, and/or translating their experience into the real world as a vote or other political action. This conceptualization of the public sphere changes it from a sociological phenomenon that can be observed and measured from the outside to a psychological or experiential phenomenon that depends on participants’ states of mind and actions. It means that even when several people are in the same virtual space (like a SNS) some may be considered to be in a public sphere while others are not.

The affordances of SNSs can facilitate the public sphere envisioned by Habermas [7,8,20] as citizens can gather together virtually, irrespective of geographic location, and engage in information exchange and rational discussion as long as they have an Internet connection. However, only when the technology is used in context does its utility emerge as a resource for political deliberation. Our study has shown that the “potential public spheres” available through social media were enacted, in turn, becoming “realized public spheres.” Many people may be engaging politically in various online public spheres without contributing content through lurking. Others may be more visible in the public spheres available online through SNSs. Whereas in the physical world we typically have awareness of those who are taking part in a discussion, SNSs enable users to pursue persistent discussions over time (as long as they are not deleted). The people who are posting information, opinions and contributing to discussions, may be making the more visible contributions, but there is an invisibility to effects of the contributions being made online, as many people can be affected by what is being posted without contributing themselves.

This conceptualization of the public sphere opens up several possibilities for research and it will be a challenge for researchers to understand how or measure when public spheres move from potential to realized. We aim to further investigate the degree to which SNSs become realized public spheres by employing ethnographic and survey methods to better understand both what is visible online, and what is invisible (going on through people’s minds).

The social network site of a politician during a campaign is a community space that is bounded in time and centered on a set of ideas, a common goal, and a defining event. Participants come to these spaces with multiple goals of their own, and this research shows that the distributions of these goals vary considerably for supporters of the different candidates and across time. The discrepancy in the distribution of some types of comment categories between communities with the same purpose is potentially meaningful to people involved in those communities. For example, community members among Whitman’s supporters might have benefitted from knowing that she was being asked an inordinate number of questions relative to her competitor and that questions were not tailing off as might be expected from a community that was successfully establishing its shared interests. We observed, for example, that many of the questions Whitman was asked initially were clarification questions about her positions. She was not as well defined within her voting community as Brown, who was well known before the gubernatorial race. However, later questions began to be more critical, centering on Whitman’s perceived unresponsiveness within her Facebook community. We conjecture that the weaker trend of turning away from engagement with the candidate over time for Whitman, and her large number of questions relative to Jerry Brown, are related to the degree that Meg Whitman’s community was not able to find common ground and establish a strong interest bond, they continued to be focused inward and maintain an information-gathering mode.

Similarly, the Brown camp might have been interested to know that their candidate was receiving many more requests than his competitor. Requests imply an assumption of agency, that the person being asked is, or will be, in a position to fulfill the request. Similarly, a comparative analysis of supportive statements...
would have revealed that Brown was receiving considerably more support than Whitman, even though there was more activity on Whitman’s site. Again we conjecture that the more clearly defined community interests allowed the group to turn their attention outwards as the election approached in the same manner that was observed in the 2008 presidential election [60,61].

Ren et al. [57] point out that choices made by designers of community and social systems influence the types of groups that can form and flourish within these systems. The Facebook comment area in 2010 was a general enough platform to support multiple goals and activities. However, Facebook and other SNSs are currently engaged in rapid redesign processes to support a myriad of different activities and user goals, often concurrently. This research suggests that focusing these efforts on artifacts, such as creating new picture-viewer or music-sharing features, is only one avenue of opportunity for designers. Another is to focus redesign efforts on support for different user goals, for example asking questions, making requests, telling stories, and expressing support. Some of these goals are general across different community contexts, whereas others might be context specific (e.g. expressing support is more likely to occur in the political arena, but asking questions will be a common activity for many groups).

The potential information to be gleaned from comparative analysis of categories between interest-based groups with the same overarching interest (electing a candidate and understanding the issues), but with opposed subordinate goals (i.e. in competition) also offers a design space for human-computer interaction. In politics, campaigns could benefit from contrasting sentiment, question, request, and other categories across the SNSs of their competitors. In contexts outside of politics, contrasting patterns could be used to gauge the health or evolutionary stage of a community’s development.

Finally, community members themselves could benefit from the design of social browsers that are sensitive to comment categories and user intentions. A new community member might, for example, be interested in browsing all of the questions that have been asked within the SNS in order to understand the issues better. Alternatively, a user might be interested in seeing all critical statements, or finding out how responsive a candidate or other members had been to requests. A “social affect browser” might be able to group comments in terms of sentiment and show users all of the angry comments or all of the supportive comments. A “social intention browser” might be able to group comments into motive categories such as questions, requests, and stories. Users would then be able to filter or navigate through different intentions. Admittedly, such judgments pose a difficult problem of linguistic analysis, but current interfaces to SNSs like Facebook, which focus mainly on chronology or friend-network structure to support browsing and navigation, are difficult to negotiate and fail to support many potential community-building activities.

On the other hand, social browsers sensitive to sentiment and intention might also pose ethical challenges for end users and designers. For example, outside of the public realm of politics, it would be possible to determine if the comments on an individual’s SNS were highly negative relative to his/her friends.

SNSs are the new public sphere for civic engagement. It is critical to understand how civic SNSs should evolve over time and to understand patterns that might indicate potential problems requiring attention. One of the core, shared interests of a civic community is the overall functional health of the community itself and its members. Information exists in the comment corpus of civic SNSs that could provide clues to the health and concerns of the community and that provide design opportunities for building and maintaining community through sentiment- and intention-based navigation, searching, and browsing.
Acknowledgments

This research was funded by NSF award No. IIS-1064852 to the first author. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References


C. Mascaro and S. Goggins, Brewing up citizen engagement: The coffee party on Facebook, Communities and Technologies 2011, Brisbane, Australia, (29 June–2 July 2011).


Z. Papacharissi, Democracy online: Civility, politeness, and the democratic potential of online political discussion groups, New Media & Society 6(2) (2004), 259–283.


T. Postmes and R. Rojas, Asymmetries in attachments to groups and to their members, Personality and Social Psychology Bulletin 20(5)


