Angry Words: A Content Analysis of Emotion in Political Blogs and Op-Ed Columns

Mark Leccese, Emerson College, USA

Meredith M. Regan, Dana-Farber Cancer Institute, USA

Abstract

Since the Internet became widely available in the 1990s, scholars have studied the online discourse around public affairs and politics. Davis (1999, p.162) studied political conversations on the now little-used Internet message board system Usenet and found what he calls "discussion dominance" in which posters "launch into rhetorical excess" and talk not to one another but past one another. Chait (2007), is his examination of the "netroots" movement, reported that bloggers avoid moderate ideas and language at all costs, seeing moderation as a concession. Johnson and Kay (2004) surveyed blog readers online and found that two-thirds wanted not fairness from blogs but bias that supports their views.

Writers in both the popular press and academic journals have argued that political blogs have increased the rancor and divisiveness in public life. Keen (2007, p. 3) writes that blogs have "perfected" political extremism. Habermas (1999) believes that rational public discourse requires sincerity in a deliberative space, but Sunstein (2007) argues that the challenge to Habermas' view is that blog readers who read primarily like-mind points of view are more extreme in their political stances. Utych (2011) found that readers of political blogs are more likely to hold extreme opinions than citizens who get their news from the traditional media. In a study of incivility in online media, Borah (2012) found that divisive rhetoric on political blogs could lead topolarization of attitudes and erode deliberative discourse.

Over the past few years, scholars have begun to focus the language and rhetoric of American political discourse online.Sobieraj and Berry (2011), continuing a strain of research into incivility in American political life, had independent coders looks for incidents of "outrage" on cable TV, talk radio, political blogs and newspaper columns using 13 categories of outrage incidents, including "name calling," "emotional display," "mockery/sarcasm," and "emotional language." One hundred percent of TV episodes and 98.8% of talk radio programs contained outrage, while "only" 82.8% of blog posts incorporated outrage writing.

Serfaty (2011) argues that political blogs thrive on strong emotions and that posts on these blogs are devised to arouse intense political emotion.

Computational Content Analysis

Farrell (2012) urges political scientists and social scientists to use computational tools on the large databases of information the Internet has generated as political discourse has moved to the Internet. One of these tools is content analysis, which Krippendorff (2004, p. 260) defines as "the drawing of inferences from text," and suggests it can be very useful in certain contexts, such as inferring the rhetorical tone of political speech.

Riabinin (2009) used computational content analysis to determine that in the 36th Canadian Parliament, words used in debate by Liberals tended to convey positive emotion, while words used by Conservatives tended to express negative affect. Riabinin calls the computational content analysis approach a "bag-of-words model," (p. 5) which is a helpful if inelegant way to describe the use of Linguistic Inquiry and Word Count Software, also known as LIWC. Gilbert, Tony and Karahalios (2009) built a sample of more than 1,000 reader's comments from 33 blogs and used Linguistic Inquiry and Word Count Software to sort comments into those that displayed positive emotion, negative emotion, anger and assent.

In trying to extend the work of Sobieraj& Berry and Serfaty using computational content analysis, this article will compare the use of words of divisiveness and anger in the work of independent political bloggers and the work of traditional media op-ed columnists. First, we need to define which bloggers this study will examine. McNair (2006, p. 120) usefully sorts "online actors" into four categories: 1) Professional-institutional actors, such as major news organizations; 2) Professional-individual actors, the few hundred political bloggers who make their living with their blogs; 3) Non-professional-institutional actors, which includes NGOs, government agencies, and campaign and lobbying groups; 4) Non-professional-individual actors, or independent bloggers. This paper will focus on the two categories of individual actors in dealing with blogs.

Our hypothesis is that the work of independent political bloggers (non-professional-individual actors) will contain more words of negative emotion, particularly anger, than

traditional newspaper op-ed writers (professional-individual actors), whose work we further hypothesize will contain more words of positive emotion.

Method

To make a comparison, five independent bloggers and ten newspaper op-ed columnists were chosen to be studied; twice the number of op-ed columnists were chosen because op-ed columnist generally published only two or three times a week while bloggers often publish several posts a day. Many blogs, including the ones chosen for this study, are written by more than one blogger but maintain a unity of voice, tone, and political outlook.

Random Selection of Samples

The highest-ranked 100 blogs were chosen from the March 2011 "Top Blogs — Politics" listing on the website Wikio News (http://www.wikio.com/blogs/top/politics), which ranked the blogs based on the number of incoming links from other blogs and the number of retweets to blogs on Twitter. The list of blogs was placed in an Excel spreadsheet and a random number was generated for each of the 100 rows. The blogs with the five lowest random-generated numbers that also had full RSS feeds were chosen for this study. They are Right Wing Watch, I Hate The Media, Daily Kos, YID With LID, and Think Progress.

To select the newspaper op-ed columnists, the most extensive survey and ranking available was consulted: *Media Matters for America's 2007 report Black and White and Re(a)d All Over: The Conservative Advantage in Syndicated Op-Ed Columns*. (http://mediamatters.org/reports/oped/) The survey contacted all 1,430 daily newspapers in U.S. in 2007 and received a response rate of 95 percent. Again, the top 100 were placed in an Excel spreadsheet and a random number was generated for each of the 100 rows. The blogs with the 10 lowest random-generated numbers that also had full RSS feeds (and were still being published) were chosen for this study. They are Steve Chapman, Linda Chavez, Suzanne Fields, Charles Krauthammer, Ruth Marcus, Deroy Murdock, Bill O'Reilly, Clarence Page, Kathleen Parker, and John Stossel.

The next step was to create a database of all the published writing of the five independent bloggers and ten newspaper op-ed columnists for full year, May 1, 2011 to April 30, 2012. The easiest way to automate this process is to collect RSS (Real Simple Syndication) feeds

into a database, but no software exists that will perform that task. The author set up Apple Mail to accept and store full-text RSS feeds from each of 15 sources and then used the Apple Automator script Mail to FileMaker Importer to move the stored full-text RSS feeds into a FileMaker Pro database. The database for each source is then exported as a tab-delimited text file and that file is converted into a Microsoft Word file. Each of the 15 files — and one file made up of all 15 databases — is now ready to be processed by the Linguistic Inquiry and Word Count software, also known as LIWC2007 (http://www.liwc.net).

At this stage, one of the op-ed columnists, Bill O'Reilly, had to be eliminated from the sample because the RSS feed of his column was changed so that a full-text feed was not allowed. The study proceeded with 9 op-ed columnists.

Linguistic Inquiry and Word Count

Tausczik and Pennebaker (2010, p. 27-28) describe how LIWC2007 searches, counts and codes words according to a dictionary compiled by the developers of the software and coded into 80 categories by judges — ranging from definite and indefinite articles to more subjective categories such as words of positive and negative emotion. The first LIWC dictionary was developed 1992-1994, the dictionary was revised 1997-2001 and revised again in 2007. The dictionary comprises almost 4,500 words and word stems. Alpets et al (2005) found that LIWC rating of words of emotion corresponded with human coding of writing excerpts. Kahn et al (2007) conducted three studies on the validity of using LIWC as a tool and concluded that LIWC is "a valid method for measuring verbal expression of emotion."

The LWIC software identifies and sorts words into three categories: Linguistic Processes (pronouns, tenses, parts of speech), Psychological Processes (social, affective, cognitive, perceptual, biological, relativity), and Personal Concerns. It can then further sort words into 70 subcategories. This study in concerned with evidence of divisiveness in the rhetoric of political commentators, so the authors chose to focus on the identification of affective psychological processes in the database texts. Tausczik and Pennebaker (2010, p. 32) state that positive emotion words "measure level of agreement" and that the use of emotion words provide "important psychological clues" to the writer's intentions and motivations (p. 37). Clearly, the subcategories "positive emotion," "negative emotion," and "negative emotion: anger" are the categories in the LIWC software that will best illuminate the level of divisiveness in the texts under study here.

Results

Between May 1, 2011 and April 30, 2012, a total of 9,356,273 words were written in 5 blogs (range: 231,322 to 3,849,584 words) and a total of 443,348 words were written by 9 newspaper op-ed columnists (range: 29,380 to 75,442 words; Table 1). Among the 14 blogs and op-ed columnists, the percentage of words in the LIWC2007 dictionaries ranged from 73.2% to 84.0%. A mean of 77.1% (SE=.28%) of bloggers' words and 78.5% (SE=1.28%) of op-ed columnists' words were in the LIWC2007 dictionaries.

As seen in Table 2, use of words in selected LIWC2007 categories ranged from 0.9% to 2.9%. Poisson regression accounting for the word count of each blog or op-ed columnist was performed to test hypotheses comparing word usage between bloggers and op-ed columnists. The study's hypothesis was not supported, in that bloggers did not use more words of negative emotion (Wald χ^2=2.30; p=.13), particularly words of anger (χ^2=.36; p=.55), more frequently than did op-ed columnists. Though the greatest difference in word usage was in the use of words of negative emotion, bloggers' use (M=2.0%, SE=.05) was observed to be less frequent than op-ed columnists' use (M=2.4%, SE=.26).

The hypothesis that newspaper op-ed columnists use more words of positive emotion also was not supported (Table 2). Op-ed columnists' use of words of positive emotion (M=2.9%, SE=.17%) was observed to be greater than bloggers' use (M=2.7%, SE=.04%), but usage was not significantly different (Wald χ^2=.80; p=.37).

Discussion

As Carroll et al (2013) have demonstrated in their studies of political polarization in the U.S. Congress dating back to Reconstruction, the U.S. House has been more polarized since 2000 than at any time since 1879, and the U.S. Senate achieved it maximum polarization since Reconstruction in 2012. The rhetoric of divisiveness in both blogs and op-ed columns may simply be a reflection of the current political climate in the United States.

Research early in the rise of independent political blogs established that these blogs create what Davis (2005, p. 124) called "discussion ghettoes," in which liberals talk only to liberals and conservatives to conservatives; the blogs serve to reinforce individual's existing views. Adamic and Glance (2005) found that bloggers link overwhelmingly to like-minded bloggers,

while Johnson and Kay (2004, p. 633) conclude that bias is seen as a virtue by readers of independent political blogs.

The finding of this study that independent political blogs do not use more divisive diction than traditional newspaper op-ed writers raises the question of whether newspaper op-ed writers today are also part of the echo chamber of partisanship. It may be edifying to conduct studies of framing, as Borah (2012) did with blogs, on op-ed columnists for traditional media outlets.This study both extends and deepens the work done by Gilbert, Tony and Karahalios (2009) and Sobieraj and Berry (2011); it uses LIWC to make a more objective measure in a sample of political blogs and op-ed columnists of what Sobieraj and Berry (2011) call "incivility," and it uses a similar method to the study by Gilbert, Tony and Karahalios (2009) but focuses on the main text of articles rather then the comments and uses a more robust database.

References

Adamic, L. and Glance, N. (2005). "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog," Blogpulse.com, March 4, 2005, Retrieved from http://www.blogpulse.com/papers/2005/AdamicGlanceBlogWWW.pdf

Alpers, G. W., Winzelberg, A. J., Classen, C., Roberts, H., Dev, P., Koopman, C., et al. (2005). Evaluation of computerized text analysis in an Internet breast cancer support group.*Computers in Human Behavior*, 21, 361-376.

Borah, P. (2012). Does It Matter Where You Read the News Story? Interaction of Incivility and News Frames in the Political Blogosphere.*Communication Research*. Published online June 14, 2012

(http://crx.sagepub.com/content/early/2012/06/14/0093650212449353)

DOI: 10.1177/0093650212449353

Carroll, R., Lewis, J., Lo, J., McCarty, N., Poole, K. and Rosenthal, H. (2013) DW-NOMINATE Scores With Bootstrapped Standard Errors [Web site with links to data]. Retrieved from http://www.voteview.com/dwnominate.asp.Updated 17 February 2013.

Chait, J. "The Left's New Machine," *The New Republic*, May 7, 2007, 18-28

Davis, R. (1999). *The Web of Politics: The Internet's Impact on the American Political System*. New York, NY: Oxford University Press.

Davis, R. (2005). *Politics Online: Blogs, Chatrooms, and Discussion Groups in American Democracy.* New York, NY: Routledge.

Farrell, H. (2012). The Consequences of the Internet for Politics.*Annual Review of Political Science*, 15(1), 35-52. doi:10.1146/annurev-polisci-030810-110815

Habermas, J. (1999). "Between facts and norms: an author's reflections." *Denver University Law Review*, 937: 940–941.

Johnson, T.J. and Kay, B.K.: "Wag the Blog: How Reliance on Traditional Media and the Internet Influence Credibility Perceptions of Weblogs Among Blog Users," *Journalism & Mass Communication Quarterly* 81, (Autumn 2004) 622-642

Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007).Measuring emotional expression with the Linguistic Inquiry and Word Count.*American Journal of Psychology*, 120, 263-286.

Keen, A. (2007). The Cult of the Amateur: How Today's Internet is Killing Our Culture. NY: Doubleday.

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology, 2nd Ed.* (2nd ed.). Thousand Oaks, CA: Sage Publications.

McNair, B. (2006). *Cultural Chaos: Journalism, news and power in a globalized world.* London and New York: Routledge.

Sunstein, C. (2007). *Republic 2.0.* Princeton, NJ: Princeton University Press.

Gilbert, E., Bergstron, T., and Karahalios, K. (2009).*Blogs Are Echo Chambers: Blogs Are Echo Chambers*. Paper presented at the 42st Hawaii International International Conference on Systems Science, Waikoloa, Big Island, HI.

Riabinin, Y. (2009). *Computational identification of ideology in text: A study of Canadian parliamentary debates.*University of Toronto, Toronto. Retrieved from www.cs.toronto.edu/compling/publications.html (MSc)

Serfaty, V. (2011). Passionate Intensity: Political Blogs and the American Journalistic Tradition. *Journal of American Studies, 45*(2), 303-316. doi: 10.1017/S0021875810001155

Sobieraj, S. a. B., J. (2011). From Incivility to Outrage: Political Discourse in Blogs, Talk Radio and Cable News. *Political Communication, 28*(1), 19-41.

Tausczik, Y. A. a. J. W. P., James W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology, 29*(1), 24-54. doi: 10.1177/0261927X09351676

Tumasjan A, S. T., Sandner PG, Welpe IM (2010). *Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment*. Paper presented at the Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Menlo Park, CA.

Utych, S. M. (2011). *Network Connections: The Differential Effects of Network News and Political Blogs on Extreme Issue Opinions*. Paper presented at the Annual Meeting of the Southern Political Science Association, New Orleans, LA.

Table 1. Word count, percentage of words in LIWC dictionary, and percentage of words in 3 LIWC word categories of 5 blogs and 9 op-ed columns during the period May 1, 2011 to April 30, 2012.

	Word Count	In Dictionary	Positive Emotion	Negative Emotion	Negative Emotion: Anger
Blog					
Right Wing Watch	904,354	78.0	2.6	2.1	1.1
Yid With Lid	877,261	78.8	3.0	2.5	1.2
I Hate the Media	231,322	74.7	2.5	2.3	1.1
Daily Kos	3,849,584	77.1	2.7	1.9	0.8
Think Progress	3,493,752	76.6	2.8	1.9	0.8
Op-ed Column					
Chapman	75,442	78.5	2.5	2.6	1.2
Chavez	36,073	79.7	2.7	2.1	0.7
Fields	29,380	79.0	2.9	2.6	1.1
Krauthammer	45,767	75.6	2.8	2.4	1.0
Marcus	59,568	77.4	2.7	2.1	0.8
Murdock	35,617	73.2	2.8	2.0	0.8
Page	72,097	78.3	3.1	2.5	1.1
Parker	58,268	81.1	3.3	2.4	1.0
Stossel	31,136	84.0	3.5	2.3	0.9

Table 2. Comparison of percentage of words in 3 LIWC word categories, for blogs and op-ed columns.

	Blogs (n=5)		Op-ed Columns (n=9)			
Word Category	**Mean ± SE**	**95% CI**	**Mean ± SE**	**95% CI**	**Chi-square**	**P-value**
Words in Dictionary	77.1% ± 0.28	(76.6%, 77.6%)	78.5% ± 1.28	(76.0%, 81.0%)	1.1	0.30
Positive Emotion	2.7% ± 0.04	(2.7%, 2.8%)	2.9% ± 0.17	(2.6%, 3.3%)	0.80	0.37
Negative Emotion	2.0% ± 0.05	(1.9%, 2.1%)	2.4% ± 0.26	(1.9%, 2.9%)	2.30	0.13
Negative Emotion: Anger	0.9% ± 0.04	(0.8%, 0.9%)	1.0% ± 0.19	(0.7%, 1.4%)	0.36	0.55