American Journal of Applied Psychology, 2014, Vol. 2, No. 5, 109-113
Available online at http://pubs.sciepub.com/ajap/2/5/2

DOI:10.12691/ajap-2-5-2

Measuring Twitter Sentiment and Implications for Social Psychological Research

Jason D. Carr*

University of Texas of the Permian Basin
*Corresponding author: jd.carr.tx@gmail.com

Received September 24, 2014; Revised October 15, 2014; Accepted October 19, 2014

Abstract This study was conducted to determine whether Twitter comments on moral issues might be classified into sentiment categories and whether underlying emotions or thoughts will influence online interactions differently than they do in the real world. A mixed methods study involving qualitative analysis was conducted to compare Twitter sentiment about a current emotionally charged topic – immigration – amongst users. Results indicated that 73% of the commenters favorably view the idea of immigration reform and/or immigrant acceptance into the U.S. and are open to online dialog. Conversely, the findings for negative comments demonstrated that users likely have underlying feelings or thoughts on the subject of immigration that in turn may cause them to interact differently online than they do in a real world setting. Stereotyping and/or bigotry may influence their communications both online and off. These findings support the need for further research to improve upon existing social psychology theories. Additionally, despite challenges present in studies of this nature, Twitter shows great promise for conducting social psychological studies in the future.

Keywords: *Twitter, immigration, sentiment, stereotypes, bigotry*

Cite This Article: Jason D. Carr, "Measuring Twitter Sentiment and Implications for Social Psychological Research." *American Journal of Materials Science and Engineering*, vol. 2, no. 5 (2014): 109-113. doi: 10.12691/ajap-2-5-2.

1. Introduction

Twitter is a rapidly growing social media platform that has an estimated 310 million visitors per month (eBizMBA, 2013). This study was conducted to determine whether Twitter comments on emotionally charged issues such as immigration might be classified into sentiment categories and whether underlying emotions or thoughts might influence online interactions differently than they do in the real world. This is important because the field of social psychology itself is focused on understanding how our thoughts or beliefs influence our interactions with others (Bremm and Kassin, 1995). By gaining a better understanding of how individuals interact online as opposed to off when discussing emotionally charged topics, underlying theories of social psychology might be improved upon in the future.

Recent years have witnessed an increasing amount of opinionated text in social media where people discuss ideas, political issues, review, and criticize objects (Chelaru, Ismail, Siersdorfer, and Nejdl, 2013). These opinions generate content that conveys community views and sentiments on various topics. Based on the sentiments expressed, it is possible to identify the social impact of messages among different societies, and how the context and framing effect of the messages shapes opinions. The public, as opposed to the crowd, is sometimes affected by factors other than face-to-face communication. The social meaning of public lies in the fact that the group may think and act in a similar fashion. Its unity is associated with an indirect psychological relationship having a common focus (Young, 1930). Latane (1981) defined social impact as influence or individual thoughts, emotions or behavior that is exerted by the real, imagined, or implied action or presence of others. The theory is associated with the magnitude of impact that an individual or group may have on one person. Research in the last few decades has revealed that how the public thinks about certain issues depends on how an issue is framed. The framing effect refers to semantically differed conceptions of the same course of action that induce preference reversals (Kahneman and Tversky, 1984). When making political choices, it is difficult to establish after the fact that the benefit of alternative course of actions for citizens will be strictly equivalent. The context in which an issue is presented to the public may also shape their opinion differently (Zaller, 1992). How citizens think about an emotionally charged issue such as immigration varies depending on whether they consider the matter in the context of concern about economic strain in the U.S., or alternatively, the well being of those entering the country to escape violence and poverty in Central America.

Historically, content analysis has made use of print literature, opinion polls and traditional surveys to derive qualitative metrics. However, in recent years, an increasing number of studies have made use of social media platforms such as Twitter to carry out content analysis. For example, in 2010, Chew and Eysenbach completed a content analysis of sentiment and public attention generated by the 2009 H1N1 influenza pandemic using Twitter. In their study, they concentrated on the public perceptions that the pandemic engendered. Their argument was that surveys, although popular, were costly, and time consuming as opposed to using a contemporary method, which they christened "infoveillance." They demonstrated that social media could be used for real-time content analysis in issues of public health, and this in turn could be useful in guiding health authorities' response to public concerns. Later, in 2011, Waters and Jamal also used Twitter to analyze how non-profit organizations communicate and build working relationships. Their study revealed that organizations are likely to use one-way communication models despite the potential for two-way dialog with users. In a significant way, this study highlighted the important role that Twitter can play in content analysis for public relations.

Conover et al. (2011) similarly investigated how Twitter and other social media shaped the public sphere and communications between different groups with different political orientations. Their content analysis concluded that a network of political retweets displayed a highly segregated and partisan structure. Their work demonstrates that Twitter can be used successfully in content analysis in the political context and that useful conclusions can be derived from such study. These studies demonstrate that Twitter can be used for content analysis across a wide array of subjects, with considerable success. In fact, Jiang et al. (2011) have proposed a number of ways to modify content analysis in Twitter so that it is target-dependent. The outcomes of each of these studies point towards useful and practical conclusions that can be derived when using Twitter for content analysis.

The most contentious issues regarding studies involving sentiment are the methodologies that have been applied to form conclusions in previous studies. The study of sentiment is particularly difficult because user comments are often subjective and must be quantified in some manner before analysis can be performed. When delving into the methodological challenges of using Twitter for sentiment measurement, a scoring model using valence and arousal should be considered. Prior to the paper published by Russell in 1980, emotional or affective states of individuals were viewed as a set of dimensions that varied independently of each other. However, Russell did an extensive study, which proposed that, rather than being independent, these emotional states were systematically related to one another. His work provided a model that has become widely accepted among psychologists, and has been used extensively in describing affective states and determining how they are interrelated. According to the model, valence refers to pleasure associated with emotion whereas arousal refers to the degree of activity induced by the emotion (Russell, 1980). In this regard, different emotional adjectives can be plotted on a one- or two-dimensional scale on hypothesized pleasure-displeasure and arousal-non-arousal dimensions. Thus, emotions are to different degrees pleasant and arousing or unpleasant and non-arousing. For example, fear is an unpleasant and highly arousing emotion while contentment is a pleasant, but non-arousing feeling. This has provided a basis for sentiment scoring that is often used in online content analysis to classify and compare emotional states of different individuals (Paltoglou, Gobron, Skowron, Thelwall & Thalmann, 2010). Affective Norms for English Words (ANEW) refers to a dictionary that provides a series of emotional ratings for approximately 1,034 English words along the three dimensions of arousal, dominance and pleasure (Bradley & Lang, 1999). Values are therefore assigned to different words that describe people's feelings, and this can be used to assign a sentiment score to each word. Ultimately, this can be used to determine a person's emotional state based on their comments. As noted by Russell (1989), this model is not without challenges. Sentiment analysis must be done within context (Wilson, Wiebe & Hoffmann, 2005). One word may be used in two opposing contexts, and this must be considered before assigning every word a score. Moreover, some words do not have definite scores to which they can be assigned.

In terms of the immigration topic itself, an emphasis placed on acculturative stress has resulted in nearly all research attempting to identify and document the negative impact of migration and corresponding acculturation via measurable indicators (Doverspike, Taylor & Arthur, 2006). Consequently, a number of studies have focused on the underlying comparative rates of maladjustment amongst young immigrants as depicted by the psychiatric or behavior disorders at the expense of normative adjustment of the children (Spielberger, 2004). Overcoming the underlying difficulties in immigration and acculturation, and to adapt readily to their surrounding at home and school, are primary emphasis of most existing immigration-related research studies. Insufficient research has been conducted which focuses on public sentiment and moral stage measurement involving U.S. citizens with regards to biases/feelings towards immigration. To date, public opinion polls remain a primary data-gathering tool and the author has found that no work with Twitter has been performed in this regard. Despite this shortcoming, many of the comments analyzed during the course of study leave little doubt that immigration is an emotionally charged topic.

With regards to stereotyping/bigotry towards immigrants, American attitudes towards them range from positive to negative, depending on the stereotype applied. Stereotypes are defined as cognitive structures that involve the perceiver's knowledge, ideas, and expectations about a given human group (Gorham, 2010, p. 94). The emotional reactions of the host society are based on the attributes related to immigrants. For instance, citizens tend to respect immigrants with competence and warmth; are contemptuous towards those who have low competence and low warmth; envy the competent but cold; and pity the incompetent but warm (Lee & Fiske, 2006, p. 765). Typically, groups that possess valuable attributes such as competence and warmth are thought to have much to offer the host country while those that are perceived as incompetent may be viewed as a drain on society. For example, Asians in the U.S., who are portrayed as model immigrants, are generally perceived as hardworking, smart and successful, while immigrants from Latin American are often stereotyped as poor, untrustworthy, criminals, and competitors for economic opportunities, (Reyna, Dobria & Wetherell, 2013, p. 342). Moreover, legal status predicts whether immigrants are perceived as part of the mainstream society or as outsiders with the lowest status, whereby those who are undocumented receive the most unfavorable stereotype (Lee & Fiske, 2006, p. 764).

2. Method

2.1. Participants and Design

During this study, individual Twitter users (N = 70) and their respective responses to a White House tweet were analyzed over a 36-hour period. The individual users each indicated that they resided in the U.S. however demographics beyond this were not readily available. A mixed methods study involving qualitative analysis was used to compare Twitter sentiment about immigration amongst users. Sentiment analysis was performed on a White House tweet (a single comment on Twitter) focused on immigration, using Twitter-Sentiment-Visualizer software. User comments were searched where tweet included both the @Whitehouse Twitter account and the term immigration. Individual comments were identified and analyzed over a 36-hour period that included these search terms. User comments were analyzed to produce an overall sentiment score using valence and arousal to define and compare individual's emotional states. Sentiment was derived through technology using machine-learning algorithms to perform analysis of text found on Twitter. A descriptive analysis of the tweets was also used to measure frequency of user comments during 3-hour increments to identify observable patterns of responses.

2.2. Measures

Individual comments were averaged and expressed as a proportion of positive comments compared with the total number of comments. Comments were graphed according to sentiment ranking based on valence and arousal ratings as well (Russell, 1980). A two-dimensional model was used to map sentiment based on a corresponding emotion. Valence and arousal values for all ANEW terms were assigned to each tweet. Mean scores for each Tweet that contained at least 2 ANEW terms were calculated using arithmetic mean. The mean valence and arousal score was computed using

$$\mu_v = \frac{\sum_{i=1}^{n} \mu_{i,v}}{n}$$

$$\mu_a = \frac{\sum_{i=1}^{n} \mu_{i,a}}{n}$$

An example of sentiment scoring is shown here.

July 10, 9:54pm: #SecureTheBorder once and for all. This should not be a **difficult** concept for #**Congress** and @WhiteHouse. #immigration @SpeakerBoehner

$v = 3.57$, $a = 6.09$

Difficult, v = [µ: 3.18, σ: 0.9] a = [µ: 5.12, σ: 2.19], fq=50

Congress, v = [µ: 4.34, σ: 1.76] a = [µ: 7, σ: 2.07], fq=50.

In the above computation, valence is calculated as 3.18 + 4.34/2 = 3.76 while arousal is calculated as 5.12 + 7/2 = 6.06. The values for each Tweet were plotted on a scale based on their overall sentiment value.

3. Results

Of the 70 individual comments analyzed, 73% of responses fell within the pleasant side of the Russell dimensional scale with high confidence ratings. The extremity on the positive side indicated that the commenter was relaxed when responding. Individuals rated on the pleasant scale were not highly aroused, but were able to control their emotions when commenting. The majority of them rated as being alert and calm when posting their comments. In as much as they were subdued, they still were able to remain calm and relaxed in the conversation on the topic of immigration. Those on the positive side of the scale were not overtly aroused by the topic. Conversely, 19 of the 70 comments fell within the unpleasant side of the scale with a lower confidence rating. The extreme range of responses on the negative side of the scale rated as having tense and nervous emotional feelings when commenting. Mean valence and arousal scores were computed and plotted for all comments to visually demonstrate sentiment of each comment (Figure 1).

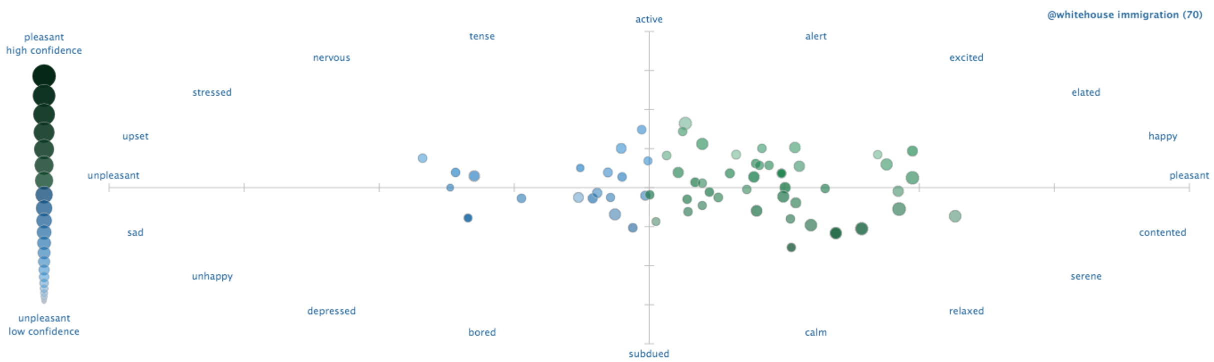

Figure 1. Mean valence and arousal scores computed and plotted to visually demonstrate sentiment of each comment

Additionally, sentiment was measured and graphed during a 36-hour period (Figure 2). Negative sentiment indicated that higher volume tweets contained stronger negative sentiment when compared to all tweets. There is strong evidence that during hours after peak volume, stronger negative sentiments were tweeted compared to hours before the peak. Furthermore, many of these tweets indicated that users held stereotypes or extreme bias towards immigrants, which may have affected both their communications and the timing of their tweets. Of positive sentiments measured, no evidence exists that higher volume hours demonstrated different positive sentiment strength compared to the lower volume hours.

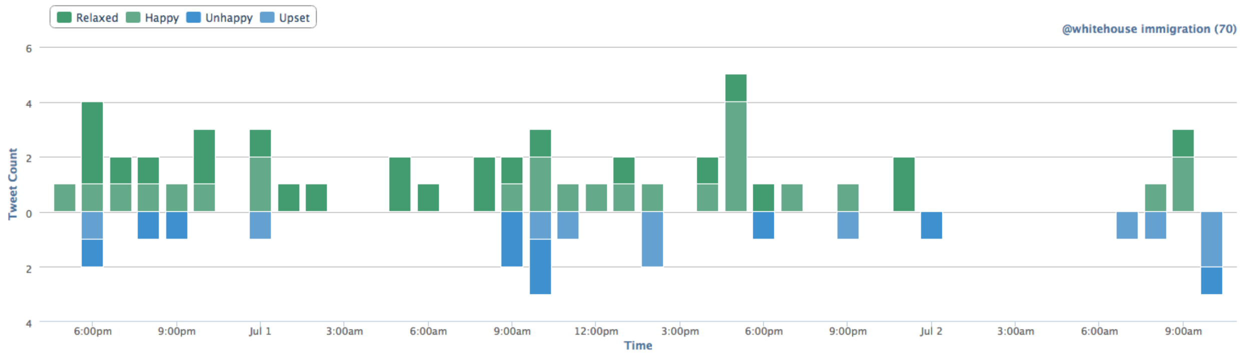

Figure 2. Sentiment measured and graphed during 36-hour period

4. Discussion

The results of this study indicate that nearly 3/4 of the users studied expressed positive sentiments toward the idea of immigration reform and/or immigrant acceptance into the U.S. Despite these findings, the author was primarily interested in those comments that fell on the negative side of the scale. Negative-rated comments analyzed were quite rigid which may indicate stronger feelings towards the subject of immigration that may not actually have been expressed on Twitter. These individuals generally did not appear to be interested in discussing the problems of immigration whereas those with positive-rated responses appeared more willing to engage in some sort of open dialog. Negative-rated comments therefore indicated that stronger underlying feelings might exist towards immigration than they were willing to express online. This indicates that these individuals may express themselves differently in a real-world setting. Based on these findings, the study successfully demonstrated that the sentiment-scoring model hypothesized by Russell is useful in content analysis, thus supporting one part of my initial hypothesis.

It is also worth nothing that several of the negative commenters appeared to hold stereotypes towards immigrants while others displayed outright bigotry based on the qualitative review of their tweets. Stereotypes provide distorted images of the targeted group and support the manner in which the groups are treated in society (Gorham, 2010, p. 98). Several of the negative commenters appeared to be misinformed on several immigration points as well. In addition, the negative commenters appeared to have little control over the stereotype impulse. The unconscious information processing that is involved pushes even low-prejudiced people to make stereotype-congruent interpretations about immigrants (Gorham, 2010, p. 101). In all negative commenter tweets examined, perceived negative qualities seemingly lead to pessimistic opinions towards immigrants. This in turn could influence their online and real-world communications. From a social psychology perspective, these findings warrant further study to examine how these types of users differ during real-world interactions as well as ways that stereotyping or biased thinking may influence their real-world communications.

5. Limitations

As mentioned earlier, content analysis using Twitter has posed several methodological challenges in previous studies. Similar weaknesses were evident in this study. For instance, although these results outlined herein appear to support my hypothesis, future research needs to be conducted using a larger number of comments for analysis. This could be accomplished by studying comments over a longer period of time, which would enable researchers to amass a much larger sample size.

Another challenge that studies such as this one face involves user demographics. For example, the 2010 Chew and Eysenbach study failed to define demographics of the population on which the study was carried out. The researchers could only guess at the percentage of American-based users in their study. The same issue was present in my study as well. Individual subject accounts were viewed and verified that each listed the U.S. as their current location, however, there is no way to actually verify whether users are truthful when setting up their accounts. One possible way to overcome this in the future is to use sentiment visualization software that maps user locations by individual IP addresses. The software used for this study didn't provide such functionality.

Previous researchers have also discovered that it is difficult, if not impossible, to determine how individuals interact with each other outside the public Twitter platform. Social interactions within Twitter do not necessarily mirror those in real life (Ovadia, 2009). Likewise, a 2011 study published by Conover et al. revealed that there was a problem in determining whether the politically polarized mood observed on Twitter was purely political or was instead instigated by other private interactions. As discussed above, comments rated as negative in this study indicated that stronger underlying feelings might exist towards immigration than was expressed on Twitter. Unfortunately, there is no feasible way to determine what these underlying feelings might be without contacting each user directly. A lack of resources during this study made this approach impossible.

6. Conclusion

Overall, this study has demonstrated that Twitter comments on moral issues can be classified into sentiment categories and that underlying emotions or thoughts might influence online interactions differently than they do in the real world when discussing an emotionally charged topic such as immigration. These findings warrant further social psychological research in this area to determine if more statistically significant findings support these conclusions. Additionally, challenges have been identified that can be taken into consideration by other researchers in the future. Finally, this study has demonstrated how social media might be used in future social psychological research efforts to improve upon existing theories. Twitter shows great promise for conducting this type of research in the field of social psychology. Based on these results, the study conducted was justified and is a positive contribution to the field of social psychology.

References

[1] Bradley, M., & Lang, P. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings* (Technical Report C-1). University of Florida, Center for Research in Phychophysiology. Retrieved from http://www.scribd.com/doc/42601042/ Affective-Norms-for-English-Words.

[2] Brehm, S. S., & Kassin, S. M. (1996). *Social Psychology* (3rd ed.). Boston, MA: Houghton Mifflin Company.

[3] Chelaru, S., Ismail, A. S., Siersdorfer, S., & Nejdl, W. (2013). Analyzing, detecting, and exploiting sentiments in web queries. Retrieved from http://www.l3s.de/~siersdorfer/ sources/2013/tweb13-sentiment-queries.pdf.

[4] Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak. *Plos One*, *5*(11). Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0014118.

[5] Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011, July). Political polarization on Twitter. *Report for the Association for the Advancement of Artificial Intelligence*. Retrieved from http://truthy.indiana.edu/site_media/pdfs/ conover_icwsm2011_polarization.pdf.

[6] Doverspike, D., Taylor, M. A., & Arthur, W. (2006). *Psychological perspective on affirmative action*. New York, NY: Novinka Books.

[7] eBizMBA (2013). Top 15 most popular social networking sites. *eBizMBA Guide*. Retrieved from http://www.ebizmba.com/articles/social-networking-websites.

[8] Gorham, B. W. (2010). Considerations of media effects: the social psychology of stereotypes: implications for media audiences. *Beyond Blackface: Africana Images In Us Media* (93-101). USA: Kendall Hunt Publishing.

[9] Jiang, L., Yu, M., Zhou, M., Liu, X., & Zhao, T. (2011). Target-dependent Twitter sentiment classification. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Vol. 1* (pp. 151-160).

[10] Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist 39*, 341-50.

[11] Kohlberg, L. (1973). The claim to moral adequacy of a highest stage of moral judgment. *The Journal of Philosophy*, *70*(18), 630-646.

[12] Latane, B., Nowak, A., & Szamrej, J. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, *97*(3), 362-376.

[13] Lee, T. L. & Fiske, S. T. (2006). Not an out-group, not yet an in-group: Immigrants in the stereotype content model. *International Journal of Intercultural Relations, 30*, 751-768.

[14] Ovadia, S. (2009). Exploring the potential of Twitter as a research tool. *Behavioral & Social Sciences Librarian*, *28*(4), 202-205.

[15] Paltoglou, G., Gobron, S., Skowron, M., Thelwall, M., & Thalmann, D. (2010). Sentiment analysis of informal textual communication in cyberspace. *Proceedings of Engage*, *2010* (pp. 13-25). Retrieved from http://www.academia.edu/3005498/Sentiment_analysis_of_ informal_textual_communication_in_cyberspace.

[16] Reyna, C., Dobria, O. & Wetherell, G. (2013). The complexity and ambivalence of immigration attitudes: ambivalent stereotypes predict conflicting attitudes toward immigration policies. *Cultural Diversity and Ethnic Minority Psychology, 19*(3), 342-356.

[17] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161.

[18] Russell, J. A., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, *57*(5), 848.

[19] Schweitzer, F., & Garcia, D. (2010). An agent-based model of collective emotions in online communities. *The European Physical Journal B: Condensed Matter and Complex Systems*, *77*(4), 533-545.

[20] Spielberger, C. (2004). *Encyclopedia of Applied Psychology*. Amsterdam, Netherlands: Elsevier.

[21] Waters, R. D., & Jamal, J. Y. (2011). Tweet, tweet, tweet: A content analysis of nonprofit organizations' Twitter updates. *Public Relations Review*, *37*(3), 321-324.

[22] Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing* (pp. 347-354).

[23] Young, K. (1930). Public opinion. In Chapter 24, *The Organs of Public Opinion. Social Psychology: An Analysis of Social Behavior* (pp. 570-598*)*. New York, NY: Alfred A. Knopf.

[24] Zaller, J. R. (1992). *The nature and origins of mass opinion*. New York, NY: Cambridge University Press.