JOURNAL OF
INFORMATION SYSTEMS APPLIED RESEARCH

In this issue:

JOURNAL OF
INFORMATION SYSTEMS APPLIED RESEARCH

Editors

Scott Hunsinger
Senior Editor
Appalachian State University

Thomas Janicki
Publisher
University of North Carolina Wilmington

2018 JISAR Editorial Board

Facebook Fanatics: A Linguistic and Sentiment Analysis of the Most "Fanned" Facebook Pages

Alan R. Peslak
arp14@psu.edu
Penn State University
Dunmore PA 18512

Abstract

With nearly 2 billion users worldwide, Facebook is the most popular social media site in the world. Despite this popularity and ubiquity, it has been lightly studied in the literature. Our manuscript examines the most popular Facebook sites (pages) in the United States dealing with society and performs a comprehensive linguistics and sentiment analysis on these sites. Using Azure machine learning for sentiment and LIWC (Linguistic Inquiry and Word Count) for linguistics, our review finds significant similarities and differences in posts on Facebook pages that have the most fans (most popular). Implications and opportunities for further research are presented.

Keywords: Sentiment analysis, Facebook, Linguistic analysis, LIWC

1. INTRODUCTION

Facebook is the most popular social media site in the world. According to Zephoria (2017) in March of 2017, there are over 1.94 billion monthly active Facebook users. This is an 18 percent increase from the prior year. Every 60 seconds on Facebook: 510,000 comments are posted, 293,000 statuses are updated, and 136,000 photos are uploaded. One in five page views in the United States occurs on Facebook.

Facebook is perhaps the most popular form of communication in the world after verbal and telephone communications. It may be the most popular form of written communications in the world.

Because of its ubiquity and popularity it is a ripe area for research and analysis. Our manuscript analyzes one area of this massive communications vehicle, popular societal Facebook pages. Our analysis is to review posts on the most popular Facebook pages through linguistic and sentiment analysis. Our measure of popularity is based on Facebook "fans". According to Chan (2009) "In the same way that profiles on Facebook help you connect with friends, Facebook Pages allow you to interact with and stay up-to-

date on your favorite public figures, organizations and businesses. When you become a fan of a Page, you are connecting with that organization or public figure and will begin seeing status updates, photos, videos and other posts from the Page. All of the posts from Pages will appear in your home page just as they would from your friends. You can get access to videos from your favorite band, chat live with your favorite celebrity, or even get a sneak peek of new products being launched by your favorite brand through Facebook Pages."

Linguistic and sentiment analysis is the review of written or verbal communications to determine specific characteristics of a communication. These characteristics can determine specific insights or meanings within a document, message, or speech that goes beyond the simple words in the communication. As a result, they can provide deeper understanding of the communicator's intent, bias, or personality thus framing the communication in a specific context and further clarify the communicator's full message. Our study reviews postings on the popular Facebook pages through linguistic and sentiment analysis.

2. LITERATURE REVIEW

Sentiment evaluation and linguistic analysis are commonplace techniques of studies in conversation analysis. The utilization of linguistic analysis and specially the use of LIWC (Linguistic and Word Count) software program for research functions has been substantial. Back, Kufner, and Egloff (2011) analyzed 11th of September communications the usage of LIWC. Cordova, Cunningham, Carlson, and Andrkowski (2001) used LIWC to research how individuals adjusted to having breast cancers. Robinson, Navea, and Ickes (2013) used LIWC analysis of college students written self-introductions to correctly calculate course performance. Bell, McCarthy, and McNamara (2012) used LIWC to analyze gender variations in linguistic styles. Sexton and Helmreich (2000) studied airline cockpit communications via LIWC to determine mistakes and overall performance. There are many other examples of the usage of LIWC inside the literature. The use of LIWC has been properly established and customary in peer-reviewed journals.

LIWC software (Pennebaker, Booth, Boyd, and Francis, 2015) is the most researched and popular linguistic analysis tool. "The way that the **L**inguistic **I**nquiry and **W**ord **C**ount program works is fairly simple. Basically, it reads a given text and counts the percentage of words that reflect different emotions, thinking styles, social concerns, and even parts of speech. Because LIWC was developed by researchers with interests in social, clinical, health, and cognitive psychology, the language categories were created to capture people's social and psychological states. …. The text analysis module then compares each word in the text against a user-defined dictionary. As described below, the dictionary identifies which words are associated with which psychologically-relevant categories." (Pennebaker Conglomerates, 2015).

Both Sentiment Analysis on Facebook posts and Linguistic analysis using LIWC have been used before in the literature. Kramer (2012) studied Facebook posts via LIWC and found that emotional status updates led to higher valence-consistent posts or posts that had more emotion. Getty et. Al (2011) studied deceased persons' Facebook profile posts via LIWC and found that Facebook served as a first stage grieving mechanism as well as maintaining a bond with the deceased. Farnadi et al. (2013) reviewed Facebook posts via LIWC to determine specific personality traits of individuals.

There has been limited study of sentiment analysis within Facebook posts. LIWC specifically does not measure sentiment. Troussas et al. (2013) examined sentiment of Facebook statuses and suggested a Naive Bayes classifier for language learning. Ortigosa et al. (2014) found studied Facebook posts to determine users' sentiment polarity with the goal of tailoring elearning systems based on students' sentiments.

One of the seminal studies in Sentiment Analysis is Sentiment Analysis and Opinion Mining by Bing Liu (2012). "Sentiment analysis, also called opinion mining, is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, " Sentiment Analysis is the review of written or other forms of communication or qualitative data to determine a quantifiable and comparable measure of some form of feeling in the communication or data.

Pang and Lee (2008) deal with sentiment polarity and degree of positivity. A basic approach is to determine whether a particular communication is positive or negative. Eguchi and Lavrenko (2006) show this by retrieving selected sentiment text. "One of the first and still most used method of sentiment analysis is keyword analysis, "

3. METHODOLOGY

On March 16, 2017, a selection and review of the most popular Facebook pages was conducted. The website and company SocialBakers was used as the source of the most popular Facebook pages in the US according to number of Fans.

The purpose of the manuscript was to examine societal trends and not mere popular entertainment or games. We decided to focus on Facebook pages dedicated to Society. In order to examine potential differences due to types of Facebook pages, an overall society review based on societal categories was performed. The categories included and selected within SocialBakers were Politics, Science, CSR or Consumer Social Responsibility, Education, NGO or Non-Governmental Associations, and Professional Associations. In total Facebook posts in these six categories were reviewed and analyzed for similarities and differences.

In order to obtain posts from these popular pages, a Facebook mining tools known as Facepager was used. "Facepager was designed for fetching public available data from Facebook, Twitter and other JSON-based API. All data is stored in a local SQLite database and may be exported to csv." Appendix Figure 1 shows an

example output screen from Facepager. Facepager was invented by Till Keyling in 2011 and is actively developed and maintained by Jakob Jünger & Till Keyling. It is free of charge and open-sourced. (Keyling, Till; Jünger, Jakob, 2013)

The most popular pages in each categories and number of fans are listed in Table 1.

Politics	Barack Obama	54,588,709.00
	Donald J. Trump	21,528,119.00
Science	NASA	18,884,690.00
	NASA Sun Science	1,380,111.00
	NASA Universe Education	1,249,453.00
CSR	Johnson and Johnson Care Inspires Care	2,991,261.00
	My Black is Beautiful	2,621,868.00
Education	Harvard University	4,916,532.00
	Make Up First School of Makeup	3,937,150.00
NGO	Causes.com	8,638,052.00
	The Animal Rescue Site	7,858,016.00
Professional Associations	American Kennel Club	3,602,312.00
	USCCA	2,016,897.00

Table 1 Most Fanned FB Pages

The most recent 500 posts were retrieved for each Facebook page and analyzed (note that the Johnson and Johnson post only had 296 posts on the page and there was an unreadable post on American Kennel Club dropping their sample to 499).

Added to the spreadsheet was a sentiment variable calculated by the publicly available sentiment calculator from Microsoft Azure Machine Learning. This variable develops an overall measure of sentiment ranging from 0 (negative) to 1 (positive) with .5 being neutral. A specific multi-decimal rating is developed e.g. .54678 from this Microsoft Excel plug-in.

We also imported the posts into LIWC (Linguistic and Word Count). LIWC software results produce 93 unique measures from each of its linguistic analyses. These measures range from parts of

speech to emotional categories to word counts. For the most part these are expressed by a percentage of total words mapping to the dictionary category of each measure. The exceptions are several relating to word counts as well as calculated emotional measures. Appendix Table 1 lists the LIWC variables used. There is also a definition of each or examples of words that meet the LIWC category. One-way ANOVA was performed to find differences among all 11 pages measured as well as between pages within a category (such as Donald Trump versus Barack Obama in politics). IBM SPSS 23.0 was used to develop the ANOVA and the results follow.

4. RESULTS

Page	Score	Category	Sig. Cat.
Donald J. Trump	65.07	Politics	0.293
Barack Obama	63.50	Politics	
NASA	66.48	Science	0.000
NASA Sun Science	60.97	Science	
NASA Universe Education	69.21	Science	
Johnson and Johnson Care Inspires Care	82.33	CSR	0.194
My Black is Beautiful	80.74	CSR	
Harvard University	65.03	Education	0.000
Make Up First School of Makeup	75.72	Education	
American Kennel Club	75.01	Prof. Assc.	0.001
USCCA	70.56	Prof. Assc.	
Total	69.59		0.000

Table 2 Sentiment as measured by Microsoft Azure

As noted the Facepager posts were analyzed via Microsoft Azure Learning sentiment analysis. The table two presents the results of the analysis. This and all scores are presented on a 1-100 scale with 0 being lowest and 100 highest. For sentiment, 1 would indicate very negative valence or sentiment, i.e. negative or bad feelings; 100 would represent very positive valence or sentiment, i.e. positive or good feelings. The tables show the sentiment score for each FB page. In addition, they show the societal category

they are attributed to and finally the significance category (Sig. Cat.) is the statistical significance of the difference between the two or more pages in that category. (This is the fourth column which is **bolded**). For example, the sentiment scores of Donald J. Trump and Barack Obama are .650746 and .634953 respectively. This is a small difference and is not statistically significant at $p < .05$. In fact p is actually **.293**, well above the .05 threshold. We can therefore conclude that there is no statistical difference between posts on Donald j. Trump's FB page and Barack Obama's FB page.

Page	Score	Category	Sig. Cat.
Donald J. Trump	75.47	Politics	.936
Barack Obama	75.32	Politics	
NASA	89.76	Science	.002
NASA Sun Science	90.04	Science	
NASA Universe Education	86.17	Science	
Johnson and Johnson Care Inspires Care	73.61	CSR	.696
My Black is Beautiful	74.48	CSR	
Harvard University	84.46	Education	.405
Make Up First School of Makeup	85.69	Education	
American Kennel Club	74.61	Prof. Assc.	0.001
USCCA	68.21	Prof. Assc.	
Total	80.04		**0.000**

Table 3 Analytic

The results of the Sentiment analysis overall present interesting results. First, the total sentiment across all selected pages and categories are all generally positive at 69 out of 100 but are significantly different at $p < .001$. The positive score is not surprising since these are pages that were "fanned" by FB users. The most positive scores were for CSR (Consumer Social Responsibility) and the two popular pages Johnson and Johnson Care Inspires Care and My Black is Beautiful were not significantly different with a p value of .194. Both scored above 80. Perhaps surprisingly the 65 and 63 sentiment scores of Trump and Obama were not significantly different. The diverse educational institutions of Harvard and Make Up First School of Makeup were significantly different with Harvard very much lower than Make Up. The fans of both the American Kennel Association and the United States Concealed Carry Association both have strong positive sentiments but AKA was significantly higher than USCCA. Finally, Science scores were somewhat lower than expected and significantly different.

The rest of the analyses all use the results of LIWC. According to (Pennebaker, Booth, Boyd, and Francis, 2015) "Analytical thinking ------ a high number reflects formal, logical, and hierarchical thinking; lower numbers reflect more informal, personal, here --- and --- now, and narrative thinking.".

Page	Score	Category	Sig. Cat.
Donald J. Trump	68.31	Politics	.875
Barack Obama	68.03	Politics	
NASA	67.08	Science	0.00
NASA Sun Science	56.08	Science	
NASA Universe Education	68.15	Science	
Johnson and Johnson Care Inspires Care	85.64	CSR	.532
My Black is Beautiful	86.73	CSR	
Harvard University	66.64	Educ	.164
Make Up First School of Makeup	64.58	Education	
American Kennel Club	76.96	Prof. Assc.	.029
USCCA	80.32	Prof. Assc.	
Total	71.14		0.000

Table 4 Clout

For the most part, all society FB pages showed high analytic content (overall averaging 73.82). Not surprisingly, the most analytic were the NASA Science posts. The least analytic was the US Concealed Carry Association at 68.21. The least analytic overall category was Professional Associations, mostly due to the USCCA. The results indicate that for society issues posts are generally formal and logical not emotional. This is perhaps a surprising result since FB is often seen as a casual and informal means of communication, as in Lofters, A. K., Slater, M. B., Nicholas Angl, E., & Leung, F.-H. (2016). This

ranking suggests that FB may provide a higher level of discourse than previously proposed.

The next measure analyzed was Clout.

"Clout ------ a high number suggests that the author is speaking from the perspective of high expertise and is confident; low Clout numbers suggest a more tentative, humble, even anxious style." (Pennebaker, Booth, Boyd, and Francis, 2015).

In general, all society FB pages have a high level of clout or level of confidence. The highest are the Consumer Social Responsibility pages. This suggests they are trusted and/or highly knowledgeable. Surprisingly, the lowest level of confidence was in Education and Science categories. This perhaps suggests the more inquisitive and exploratory nature related to these fields. It is interesting to note that posts on both Trump and Obama Politics pages are about average for these categories and are not significantly different in clout. The USCCA has the highest confidence after the CSR pages.

Page	Score	Category	Sig. Cat.
Donald J. Trump	37.45	Politics	.001
Barack Obama	30.11	Politics	
NASA	47.99	Science	.015
NASA Sun Science	50.62	Science	
NASA Universe Education	44.96	Science	
Johnson and Johnson Care Inspires Care	27.25	CSR	.243
My Black is Beautiful	29.96	CSR	
Harvard University	28.94	Education	0.000
Make Up First School of Makeup	15.46	Education	
American Kennel Club	21.60	Prof. Assc.	0.000
USCCA	28.77	Prof. Assc.	
Total	33.23		0.000

Table 5 Authentic

The authenticity (Authentic) measure averages only 30.44. "Higher numbers are associated with a more honest, personal, and disclosing text;

lower numbers suggest a more guarded, distanced form of discourse." (Pennebaker, Booth, Boyd, and Francis, 2015). Our results suggest a less personal form of disclosure and more distanced.

LIWC Analytic Measures for Selected Media (Pennebaker, J., Boyd, R., Jordan, K., and Blackburn, K. (2015))

Natural speech has been analyzed to be 61 on the 0-100 scale. (See Appendix Table 2 for LIWC ratings for common forms of communication). The Societal FB posts average only 30. This proposes that FB posts are much more guarded than natural speech. This may be due to the broad public nature of FB. People may be much more leery of posting comments on FB than natural speaking. This has broad implications for analyzing the content of FB posts. These posts may not reflect all true feelings. Due to a less controversial area, the Science posts are much less guarded.

Page	Score	Category	Sig. Cat.
Donald J. Trump	58.57	Politics	.024
Barack Obama	52.87	Politics	
NASA	43.75	Science	0.000
NASA Sun Science	52.61	Science	
NASA Universe Education	55.15	Science	
Johnson and Johnson Care Inspires Care	84.80	CSR	.020
My Black is Beautiful	79.48	CSR	
Harvard University	51.68	Education	.228
Make Up First School of Makeup	48.93	Education	
American Kennel Club	71.58	Prof. Assc.	0.000
USCCA	54.46	Prof. Assc.	
Total	58.47		0.000

Table 6 Tone

Tone differs somewhat from sentiment. "Emotional tone ------ a high number is associated with a more positive, upbeat style; a low number reveals greater anxiety, sadness, or hostility. A number around 50 suggests either a lack of

emotionality or different levels of ambivalence." The emphasis here is on style, upbeat or sad. Overall the average was slightly more than ambivalent. Trump posts were significantly more upbeat than Obama. The CSR pages were very upbeat as was the American Kennel Club. NASA as a whole was a bit more downbeat than the NASA specific sites. The rest hovered around the non-emotional average. None of these results seem surprising and are consistent with their messages and followers.

Page	Score	Category	Sig. Cat.
Donald J. Trump	8.29	Politics	.000
Barack Obama	6.54	Politics	
NASA	5.36	Science	0.000
NASA Sun Science	5.02	Science	
NASA Universe Education	8.87	Science	
Johnson and Johnson Care Inspires Care	11.37	CSR	.085
My Black is Beautiful	12.40	CSR	
Harvard University	5.69	Education	0.000
Make Up First School of Makeup	4.16	Education	
American Kennel Club	8.78	Prof. Assc.	0.000
USCCA	11.64	Prof. Assc.	
Total	7.88		0.000

Table 7 Pronoun

A high percentage of pronouns reveals a more personal and informal style. Natural speech rates at 15% usage of pronouns. All our Facebook Societal posts are less than this, averaging only 7.97. The most personal were in CSR posts at between 11 and 12. The lowest were Science and Education at 5. In general, it can be said that Societal FB page posts reveal little of personal style. This is consistent with the authenticity measure which showed more guarded communications.

Page	Score	Category	Sig. Cat.
Trump	5.96	Politics	0.011
Barack Obama	4.79	Politics	
NASA	1.50	Science	0.000
NASA Sun Science	2.49	Science	
NASA Universe Education	2.54	Science	
Johnson and Johnson Care Inspires Care	9.83	CSR	0.000
My Black is Beautiful	7.41	CSR	
Harvard University	3.59	Education	0.000
Make Up First School of Makeup	2.34	Education	
American Kennel Club	7.86	Prof. Assc.	0.000
USCCA	4.19	Prof. Assc.	
Total	4.58		0.000

Table 8 Positive emotions

Page	Score	Category	Sig. Cat.
Trump	1.43	Politics	0.798
Barack Obama	1.48	Politics	
NASA	0.38	Science	0.535
NASA Sun Science	0.32	Science	
NASA Universe Education	0.34	Science	
Johnson and Johnson Care Inspires Care	0.13	CSR	0.045
My Black is Beautiful	0.32	CSR	
Harvard University	0.93	Education	0.000
Make Up First School of Makeup	0.26	Education	
American Kennel Club	0.66	Prof. Assc.	0.000
USCCA	1.58	Prof. Assc.	
Total	.73		0.000

Table 9 Negative emotions

Natural speech tends to average about 5% positive emotions and 1% negative emotions. The results of the FB post analysis suggests that FB posts for the most popular "Fanned" FB pages mirrors this general level and ratio. The actual results are 4.55 positive emotions and .81 negative emotions. Thus, generally, emotionally we communicate similarly on FB fan pages and in natural speech. If confirmed through further studies this is an important finding with regard to online versus in-person communications.

Also there are very interesting results with regard to individual pages and categories. In Politics, Donald J. Trump fans have statistically significant higher positive emotions than Barack Obama, but show no difference in negative emotions. Science also has significant differences in positive emotions but not in negative emotions. All the Science pages show little positive emotions though, with NASA almost 0 at 1.5%. CSR pages have much higher positive emotions and much lower negative emotions. There are significant differences in Johnson and My Black however. Johnson has higher positive and lower negative.

Page	Category	Ratio	Average
Donald J. Trump	**Politics**	4.16	3.702
Barack Obama	Politics	3.23	
NASA	**Science**	3.94	6.39
NASA Sun Science	Science	7.78	
NASA Universe Education	Science	7.47	
Johnson and Johnson Care Inspires Care	**CSR**	75.61	49.38
My Black is Beautiful	CSR	23.15	
Harvard University	**Education**	3.86	6.43
Make Up First School of Makeup	Education	9	
American Kennel Club	**Prof. Assc.**	11.90	7.28
USCCA	Prof. Assc.	2.65	
Total		5.61	

Table 10 Positive to negative emotions ratio

Education has significantly higher positive emotions for Harvard but also higher negative than Make Up. Finally, the American Kennel Club has significantly higher positive emotions and lower negative emotions than US Concealed Carry.

5. FURTHER DISCUSSION

Overall, the sentiment and linguistic analyses of popular FB pages yields some very stimulating results. The sentiment measure for Facebook pages that individuals have fanned are generally positive. Scores ranged from 65 to 82 with Politics on the lower end and Consumer Social Responsibility on the higher end. This sentiment analysis is somewhat confirmed by the positive and negative emotion ratings that LIWC has developed. The highest positive emotions were for CSR but the lowest positive emotions are for the Science category. Further analysis of the ratio of positive to negative emotions however better supports overall sentiment.

Politics actually had the lowest ratio of positive to negative emotions of any category at 3.70. Science was actually above total at 6.39. CSR had an incredibly high positive to negative ratio of 49.4.

Comparing this ratio to common other forms of communication also yields interesting results.

	Positive emotions	Negative emotions	Ratio
Blogs	3.66	2.06	1.78
Expressive writing	2.57	2.12	1.21
Novels	2.67	2.08	1.28
Natural Speech	5.31	1.19	4.46
NY Times	2.32	1.45	1.60
Twitter	5.48	2.14	2.56

Table 11 Other Communications Positive to negative emotions ratio

For all Facebook posts we analyzed the ratio of positive to negative emotions compared to Blogs, expressive writing, novels, natural speech, NY Times, and Twitter. The FB posts are significantly more positive in emotion than all these forms of communications. It is posited then that these FB posts do not accurately reflect how we communicate in everyday life or other forms of communication. Rather they are artificially positive in their message and content and provide

support and good feelings not consistent with other forms of discourse.

All the FB pages studied scored high in analytic measures ranging from 90 for NASA to 68 for USCCA. These measures were over all other forms of communication in Appendix Table 2 except the New York Times. Posts were generally formal and logical. In general, confidence levels of the reviewed FB posts were confident and strong. Scores ranged from a high of 80 for USCCA to 56 for NASA Sun Science. By comparison, regular speech only has a 52 Clout measure and the New York Times is only 68. Coincidentally, the Clout scores for Donald J. Trump and Barack Obama were also 68. Level of authenticity for all FB posts was low; much lower than natural speech, novels, or blogs (61, 75, 60 respectively). The FB posts ranged from a low of 15 for Make Up School to a high of 50 for NASA Sun Science. All others were below 50. This suggests a level of guardedness in what is posted online. Many are near the guardedness of the New York Times which is at 25. Overall Tone of the posts for the most part was somewhat neutral, though the CSR posts were very upbeat at 80 and above. Interestingly, natural speech occurs at about this 80 level, much higher than other forms of communication such as blogs (55), the New York Times (44) and novels (37) as well as most of our studied FB posts (average 58). The personal aspect of our FB posts as measured by the use of pronouns was for the most part much less than natural speech (15) or the New York Times (21). Generally it can be said that FB posts are less personal than most other forms of communication.

6. CONCLUSION

It should be noted that there are limitations to the study. First, only one day was selected but many posts were prior to that day. Further duplication studies should be performed over time. Next, only 500 posts were used for each analysis. Greater numbers may yield different results. Finally only top "fanned" pages were used. Different results may be obtained by using lesser popular pages. Despite these limitations, this study has demonstrated a series of important results. This study of sentiment analysis extends the work of many applied IS research including highly cited works from Computers in Human Behavior, Communications of the ACM, SIGCHI conferences and Expert Systems with applications, *Foundations and Trends® in Information Retrieval.* First it defines, presents and demonstrates an example and interpretation of linguistic analysis and sentiment analysis using

one of the most researched and developed tools, LIWC. Researchers and practitioners can use this manuscript as a source and guide for developing their own linguistic analysis of any communication. Second, the study illustrates the results of Facebook posts metrics as they compare to other forms of computer-mediated communications. Researchers and practitioners can reliably use this comparison for other forms of computer-mediated communications. Finally, the study analyzes Facebook posts via linguistic and sentiment of the most popular FB Society sites and categories. The results show significant differences in all areas of sentiment and linguistic analyses. There are also significant differences within categories. Researchers can use these findings to compare and contrast Facebook posts to their linguistic characteristics. Societal social network Facebook page hosts can use these findings to improve their overall sentiment and linguistic metrics if they choose.

7. REFERENCES

Back, M. D., Küfner, A. C., & Egloff, B. (2011). Automatic or the people? Anger on September 11, 2001, and lessons learned for the analysis of large digital data sets. *Psychological Science*, *22*(6), 837-838.

Bell, C. M., McCarthy, P. M., & McNamara, D. S. (2012). Using LIWC and Coh-Metrix to investigate gender differences in linguistic styles. *Applied Natural Language Processing: Identification, Investigation, and Resolution, Information Science Reference, Hershey, PA*, 545-556.

Chan, K. (2009) Facebook Tips: How do I find and "fan" a Page? https://www.facebook.com/notes/facebook/facebook-tips-how-do-i-find-and-fan-a-page/185405397130/

Cordova, M. J., Cunningham, L. L., Carlson, C. R., & Andrykowski, M. A. (2001). Social constraints, cognitive processing, and adjustment to breast cancer. *Journal of consulting and clinical psychology*, *69*(4), 706.

Eguchi and V. Lavrenko, "Sentiment retrieval using generative models,"in Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 345–354, 2006

Farnadi, G., Zoghbi, S., Moens, M. F., & De Cock, M. (2013, January). Recognising personality

traits using facebook status updates. In *Proceedings of the workshop on computational personality recognition (WCPR13) at the 7th international AAAI conference on weblogs and social media (ICWSM13)*. AAAI.

Getty, E., Cobb, J., Gabeler, M., Nelson, C., Weng, E., & Hancock, J. (2011, May). I said your name in an empty room: grieving and continuing bonds on facebook. In *Proceedings of the SIGCHI Conference on human factors in computing systems* (pp. 997-1000). ACM.

Keyling, Till; Jünger, Jakob (2013). Facepager (Version, f.e. 3.3). An application for generic data retrieval through APIs. Source code available
from https://github.com/strohne/Facepager.

Kramer, A. D. (2012, May). The spread of emotion via Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 767-770). ACM.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1-167.

Lofters, A. K., Slater, M. B., Nicholas Angl, E., & Leung, F.-H. (2016). Facebook as a tool for communication, collaboration, and informal knowledge exchange among members of a multisite family health team. *Journal of Multidisciplinary Healthcare*, *9*, 29–34. http://doi.org/10.2147/JMDH.S94676

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, *31*, 527-541.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, *2*(1-2), 1-135.

Pennebaker Conglomerates (2015) *LIWC How It Works*. http://liwc.wpengine.com/how-it-works/

Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015*). Linguistic Inquiry and Word Count: LIWC2015*. Austin, TX : Pennebaker Conglomerates (www.LIWC. net).,

Pennebaker, J., Boyd, R., Jordan, K., and Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX. University of Texas at Austin.

Robinson, R. L., Navea, R., & Ickes, W. (2013). Predicting final course performance from students' written self-introductions: A LIWC analysis. *Journal of Language and Social Psychology*, 0261927X13476869.

Sexton, J. B., & Helmreich, R. L. (2000). Analyzing cockpit communications: the links between language, performance, error, and workload. *Journal of Human Performance in Extreme Environments*, *5*(1), 6.

SocialBakers. (2017) Top Facebook pageshttps://www.socialbakers.com/statistics/facebook/pages/total/united-states/society/science/

Troussas, C., Virvou, M., Espinosa, K. J., Llaguno, K., & Caro, J. (2013, July). Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning. In *Information, Intelligence, Systems and Applications (IISA), 2013 Fourth International Conference on* (pp. 1-6). IEEE.

Zephoria (2017) https://zephoria.com/top-15-valuable-facebook-statistics/

Appendices

**Appendix Figure 1 Sample Facepager outpu
t**
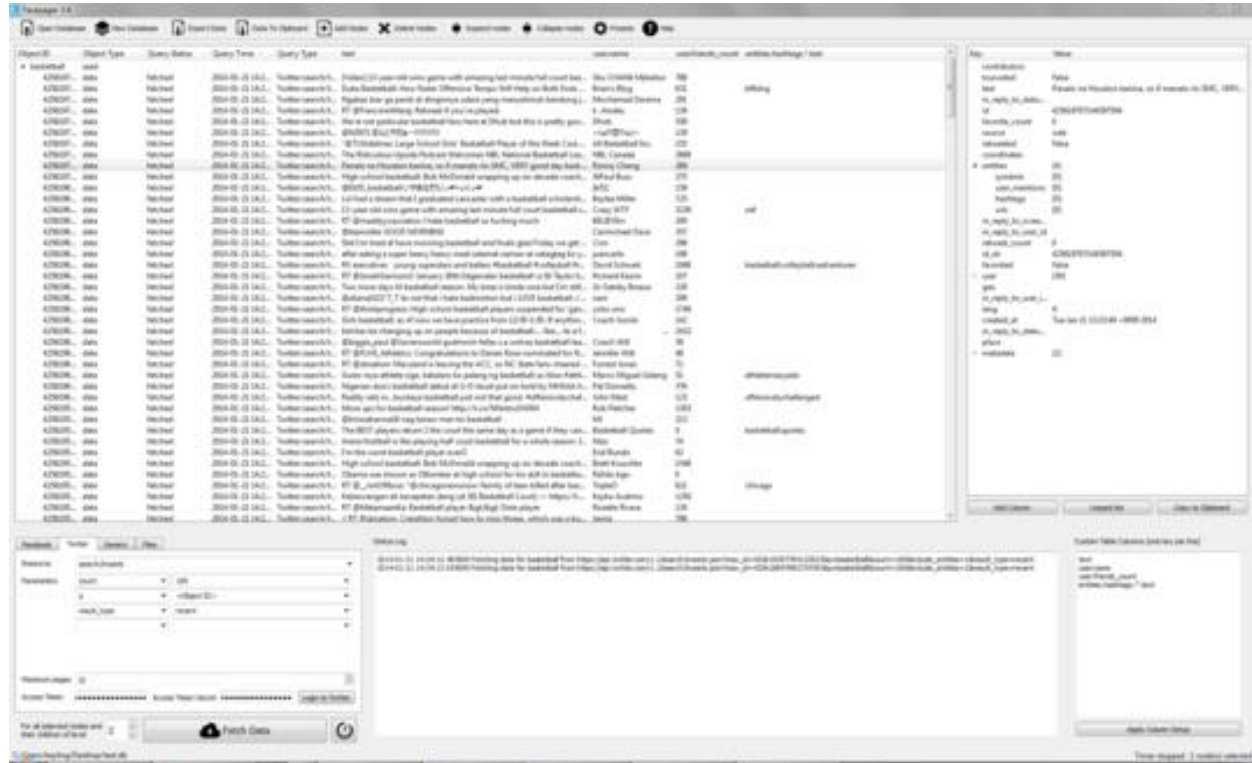
Appendix Table 1 Dependent and Independent variable table

Variable	Meaning or examples
SentN	Overall sentiment of tweet (0 negative to 1 positive)
Analytic	reflects logical thinking versus narrative
Clout	Confident (high score) versus Tentative (low score)
Authentic	Honest versus Guarded
Tone	Upbeat versus sad
Dic	Number of words in LIWC dictionary (suggests ease of read)
Pronoun	I, them, itself
Posemo	love, nice, sweet
Negemo	hurt, ugly, nasty

Appendix Table 2. LIWC Analytic Measures for Selected Media (Pennebaker, J., Boyd, R., Jordan, K., and Blackburn, K. (2015))

	Blogs	Expressive writing	Novels	Natural Speech	NY Times	Twitter	Explanation
Analytic	49.89	44.88	70.33	18.43	92.57	61.94	Logical versus informal
Clout	47.87	37.02	75.37	52.27	68.17	63.02	Confident versus humble
Authentic	60.93	76.01	21.56	61.32	24.84	50.39	Honest versus guarded
Tone	54.5	38.6	37.06	79.29	43.51	72.24	Upbeat versus hostile
Dictionary	85.79	91.93	84.52	91.6	74.62	82.6	Nontechnical, Number of words in LIWC dictionary (suggests ease of read)
Pronoun	16.2	16.2	18.03	15.15	20.92	7.41	Personal and informal
Positive emotions	3.66	2.57	2.67	5.31	2.32	5.48	Happy
Negative emotions	2.06	2.12	2.08	1.19	1.45	2.14	Sad or angry