9-2014

A Crowdsourcing Approach to Identify Common Method Bias and Self-Representation

Margeret A. Hall
University of Nebraska at Omaha, mahall@unomaha.edu

Simon Caton
National College of Ireland

Online-gathered social data raises challenges for researchers aiming to unobtrusively apply publically accessible online data to generalizable social models. The trove of potential data is vast, but the ability of researchers to verify its authenticity is low. Pertinent questions on the measurement of social indicators are: the verification of data gained online (e.g., controlling for self-representation on social networks), and appropriate uses in community management and policy-making. Across platforms like Facebook, LinkedIn, Twitter, and blogging services, users (sub)consciously represent themselves in a way which is appropriate for their intended audience (Qiu et al. 2012; Zhao et al. 2008). However, scholars have not yet adequately addressed controlling for self-representation in online social networks, or the propensity to display socially responding characteristics or censor oneself in online fora, in their analyses (Zhao et al. 2008; Das & Kramer 2013). As such, researchers on these platforms risk working with 'gamified', socially responding, or online disinhibitive (trolls) personas that goes beyond efforts to contain Common Method Biases (CMB) (Linville, 1985; Suler, 2004; Podsakoff et al., 2003). What has not been approached in a systematic way is the verification of such data on offline and actual personality. This leaves the open question of alignment of unobtrusively gathered data and online self-reported data. In this paper, we focus on the alignment of traditional survey methods with unobtrusive methods to gather profile data from online social media via crowdsourcing platforms.

This research has two aims: 1. Establishing the relationship between offline and online personalities via survey responses and self-produced text, and 2. Mitigation of biases in both traditional survey methods and in publically sourced data. In response to these research aims, we hypothesize that self-representation can be identified, and thus eventually be controlled for in broad social models. Surveys are prone to rater and item effects (Podsakoff et al. 2003) and online data is susceptible to context effects. By first addressing and mitigating CMB, and then using this to isolate further personality enhancements and/or misrepresentations, we create a research method for the analysis and use of publically gathered data with minimal bias. The proposed alignment method attempts to address concerns of CMB in analysis of publically sourced data by investigating and addressing common rater effects, item characteristics effects, item context effects, and measurement context effects (Podsakoff et al. 2003) in a multi-method approach. Our research creates an estimation function to lessen the distortion created by CMB and self-representation.

For this, the study employed the popular crowd work platform Amazon Mechanical Turk, receiving survey responses and Facebook Timeline data from 509 workers. The results of this experiment are structured in the following way: section two reviews related literature, section three discusses the methodology of the Human Intelligence Task (HIT), computational mechanisms behind the study's tool, and employed statistical methods. Sections four and five present and discuss the results. Section six summarizes the paper's contribution, limitations, and points out areas for future work.

Related Literature

It is indisputable that social media and the Internet more broadly reshaped information disbursement and processing. This leads to specific challenges for the policy arena in adapting to the management of communication. As a generalization, social media users can be split into two groups: users who search for information, and users who participate in information formation (Kushin & Yamamoto 2010; Auer 2011). Especially important for policy professionals is observing and managing the effects of information creators on information recipients (Auer 2011), as this has been proven to contribute to what is known as the 'spiral of silence' in public opinion, both on and offline (Noelle-Neumann 1974). In (Hampton et al. 2014), it was found that social media users are even less likely to express their opinions offline if they believe they differ from the majority opinion. This need becomes ever more pressing in the face of recent findings from Pew Research, that 30% of Americans receive their news from Facebook, 10% from YouTube, and 8% from Twitter.[1] Especially that oftentimes users actively search for opinions that mirror their own, the authenticity of citizen-produced news and its reliability is of up most importance. We note that this could also be a contributing factor to the findings of (Kramer et al. 2014). Self-representation has been discussed in several works for online and offline fora. These studies discuss that one's tendency to disclose information about personal experiences emanate from an associated intrinsic value. While many methods including surveys and interviews, and ethnographic research can identify self-representation, text analytics is a promising research design for the identification and mitigation of self-representation bias in data.

Self-representation, emotional disclosure, and online social networks

Self-representation in real life works with verbal and non-verbal impressions, whereas the entire non-verbal component is omitted in online social networks. Users compensate by means of text, music, video and pictures. However, these expressions do not replace personal expressions and gestures adequately. Moreover, in real life direct communication is often the social norm (Hoever, 2010).

Presentation of self in terms of online media was theoretically addressed by (Hogan 2010). He contends self-representation is an increasingly frequent strategy in online participation. Following (Goffman 1959), Hogan addresses digital ‘exhibitions’ and ‘curators’ where exhibitions are defined as status updates, listicles, or photos and the virtual curator creates the digital contents. In setting the terms of self-presentation in theatrical terms, this work makes distinct that self-representation is the display of the ideal self, rather than a pattern of deception. The value of self-representation is supported in the findings of (Mehra et al. 2001), who look at high and low self-monitoring by employees in a high technology firm in order to disentangle if there are effects on profits. They find that high self-monitors

[1] A summary of the Pew report can be found here: http://www.journalism.org/2014/03/26/8-key-takeaways-about-social-media-and-news/

are more likely to occupy preferential positions and have higher social network density than low self-monitors.

In social networks communication is more indirect. Status updates, uploading pictures or inserting information in the "About Me" section is not directed to anyone specifically. Although one approximately knows who may be reached, they do not know who will respond. Facebook is largely based on off-line relationships (Zhao et al., 2008). Since Facebook is not anonymous, the freedom of identity construction is significantly restricted. Most people use Facebook to stay in touch with real-life people, so they cannot completely detach their true identity. Users try to present a socially aspired self-image to be ‘popular’(Utz et al. 2012). It was also found that users want to make themselves seems more interesting and therefore shorten self-descriptions (Utz et al. 2012). Self-representation in real life it is also linked to time and place. One must immediately respond to an interlocutor or opponent, and this person is usually known. In social networks, one has the option to not act immediately. Local binding is eliminated with social networks (Hogan 2010; Goffman 1959).

In a study by (Ellison et al. 2006) the researchers considered an online dating environment in order to determine how honestly users depict themselves. Results of their interviews (n=34) indicate that the users who are more ‘honest’ in self-presentation have more success in dating. Nevertheless, all interviewees noted that in their online dating profiles they attempt to reveal themselves particularly positively, and have the same impression of the profiles of other users. Another notable contribution of this paper is its use of a social shaping perspective of technology, which acknowledges “the ways in which information and communication technologies (ICTs) both shape and are shaped by social practices” (Ellison et al. 2006, p 417), as opposed to a technologically deterministic perspective that focuses on the characteristics of the technologies themselves, or a socially deterministic approach that concentrates on user behaviour.

Facebook Research's study on self-censorship, the typing then deleting or posting of statuses and comments, from 3.9 million Facebook users also looks at how users alter their statements in quasi-public fora (Das & Kramer 2013). This studied has shown that 71% of users self-censor at least in some way. Male users censor more than female, and Facebook posts are more frequently regulated than comments. They find that those with higher boundaries (estimated by the amount of regulations in place on the audience of the posting person) self-censor more, and theorize that the lack of control over an audience drives self-censorship.

Emotional disclosure on Facebook, and other social networks, is different from the feeling of disclosure in real life (Qiu et al., 2012). On Facebook, feelings are shown through pictures, status updates and comments. In real life, however, a person's feelings can often be guessed through facial expressions and body posture. Studies show that self-disclosure online is generally more empathized than in real life. In (Qiu et al. 2012), it was discovered that users communicate their positive emotions more frequently via social posturing. They find that negative emotions in Facebook are hardly communicated.

Text analytics

Function words comprise approximately 55% of a given language (Tausczik & Pennebaker 2010). Function words can detect emotional states, spot when people are lying, predict where they rank in social hierarchies and the quality of their relationships, along with their Five Factor Personality Model scores and happiness levels (Pennebaker et al. 2003; Yarkoni 2010). Function words can detect emotional states; predict where people rank in social hierarchies and the quality of their relationships. Across studies when participants were lying they used more negative emotion, more motion words (arrive, car, go), fewer exclusion words, and less first-person singular (Chung & Pennebaker, 2014). People who are being interrogated, for example, use far more I-words if telling the truth than if lying. Deceptive statements are balanced in descriptiveness: enough description is required to convince the other person of an untruthful statement but too much information might reveal inaccuracies (Chung & Pennebaker 2014).

Different patterns of function words reveal important parts of people's personalities. Formal writers tend to be more concerned with status and power and are less self-reflective. They drink and smoke less and are mentally healthy, but tend to be less honest. As people age, their writing styles tend to become more formal (Pennebaker et al. 2003). Analytical writers make distinctions; they attain higher grades, tend to be more honest, and are more open to new experiences. They also read more and have more complex views of themselves (Pennebaker et al., 2003).

The Linguistic Inquiry and Word Count (LIWC) tool has robust, reliable results in measuring personality (Pennebaker et al. 2007; Yarkoni 2010). Its multi-tiered dictionary allows for a flexible set of covariates on which to align personality and happiness with the text-based data. LIWC was not intended to be used on short informal text, but to analyse text of expressive and therapeutic writing sessions usually containing more content than the average tweet (about fifteen words) or Facebook update (about nine words) (Kramer 2012; Kramer 2010). However, its expansive psychometric dictionary offers a unique opportunity to reveal the latent emotional context of text-based data. Per cent based information shows not only the latent context but gives the researcher a mechanism by which to see the relative worth of categories in speech.

Using short informal text as a foundation of public sentiment measurement differs from other text due to the shortness of the text and the different language used (Thelwall et al. 2010). Short text like tweets and status updates, with their small average word count, forces the writer to get quickly to the point. Yet, word count restrictions foster the usage of abbreviations and emotional tokens, leading to informal text containing not only slang but also abbreviations in different forms, emoticons in various forms and styles as well as truncated sentences (Wang et al. 2014).

Methodology

To facilitate our study, 509 Amazon Mechanical Turk workers completed psychometric surveys (John et al. 1991; Huppert & So 2013) and questions on Facebook usage (Ewig 2011) via a Facebook application, from which 469 wholly-recorded questionnaires were returned. The list of questions is available in Appendix I. The works (Hall, Caton, et al. 2013; Hall, Glanz, et al. 2013) show that the Big-Five Inventory and Human Flourishing are reliably recorded in an online environment. An initial screening question based on reading attentiveness was employed in order to minimize 'click-through' behaviour (Berinsky et al. 2012). Due to the question structure and number of question, nine minutes was established as the minimum amount of time needed for completion, and workers who completed in less than nine minutes were excluded from the analysis. The study was launched over a 24-hour period to accommodate differences in time zones.

A summarized privacy statement was presented on the entry page of the HIT, and full privacy statement was available on request, detailing the uses of data and steps undertaken to guarantee participant privacy. As participants completed the survey, a PHP-based Facebook application simultaneously accessed their unique Facebook ID, and via Facebook's Open Graph API (application programming interface) accessed participants' Facebook timelines (Figure 1). Payments of US$ 0.74 were issued at the end of the survey, which equates to 1 cent per question. Participants' IDs were one-way hashed, with profile, survey, and worker payment being tied to the hashed ID. As the data is stored to disk, the hashing of IDs is necessary to maintain user anonymity.

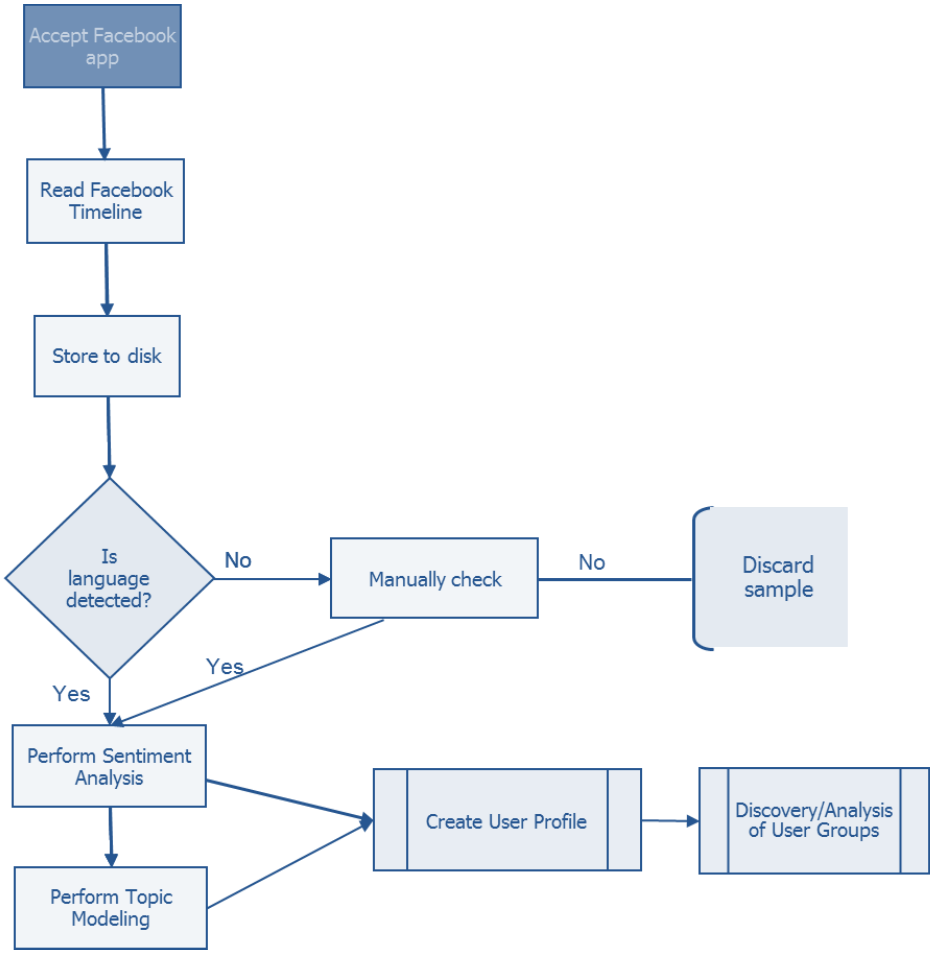

Figure 1. Workflow illustrating the steps to acquire, analyse, and interpret text data.

Workers were given an option to opt out of the HIT at the stage where it linked to their Facebook profile or abandon the HIT at any other point. Privacy-aware users were able to hide their activities from the app. Regardless of a users' privacy settings allowing timeline extraction or not, workers were paid with survey completion. The app extracted only posts, i.e., status updates, participants made to their timelines. Other post types such as shares, profile updates, etc. are excluded as they are not fully self-produced texts. This type of constraint can create first-order bias by potentially culling messages from the list of retrieved posts. However, as we are considering presentation of the self, and mitigation of possible bias in self-presentation, comments from other users are not immediately helpful. As this study is not a network study, second order bias is not considered here (González-Bailón et al. 2014).

To analyse participant timelines offline, we parsed the JSON object retrieved from Facebook storing them in flat files so that they could be imported into LIWC for sentiment retrieval. Procured data is stored initially in JSON objects (one per participant) and represents the entire timeline and basic information – this format mimics the Facebook representation of data, only without pagination. To analyse Facebook data, we first partition the data set with various granularities, i.e., per hashed ID or ID groups, and then temporally i.e., weekly, monthly, or the complete collection of posts for the entirety of the timeline. Compiling the data in this manner allows us to execute studies with LIWC at multiple granularities and time samples. Currently, the LIWC analysis is performed manually as LIWC does not facilitate automated invocation, preventing a post-by-post analysis due to the volume of calls needed.

Two statistical procedures are heavily utilized in this work, namely Spearman's ρ and automatic Linear modelling (SPSS version 22). While linear relationships exist in the data, some cases are non-normally distributed. (Fowler 1987) notes that Spearman's ρ outperforms other correlation methods in cases of contaminated normal distributions, and is robust to Type III errors (correctly rejecting the null hypothesis for the wrong reason(s)). This justifies the use of ρ rather than Pearson's r, in spite of the fact r tests on true values rather than ranks (thus monotonic relationships). Spearman's ρ is calculated as:

$$\rho = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

For a sample of size *n*, with the n raw scores X_i, Y_i, are converted to ranks x_i, y_i, where $d_i = x_i - y_i$, is the difference between ranks. Automatic linear modelling is employed for its facilities in automatic data preparation and handling. Regression in SPSS version 22 is ruled out as it is limited to step-wise methods only, cannot conduct an all-possible subset analysis (which is necessary here for exploratory reasons), and does not automatically identify and handle outliers. Automatic linear modelling is more robust against Type I and II errors in comparison, and can improve predictions by conducting a model ensemble (Yang 2013). Our analysis utilises the boosted, best-subset model consistent with data mining approaches. SPSS 22 defines multiple imputation general linear regression as (IBM 2011b; IBM 2011a)

$$y_i = x'_i\beta + e_i \text{ with } e_i \sim N\left(0, \frac{\sigma^2}{w_t}\right)$$

$$\text{Prior: } Pr(\beta, \log\sigma^2) \propto 1, \text{ or equivalently } Pr(\beta, \sigma^2) \propto 1/\sigma^2$$

Using the complete cases, to fit the regression model. The assumption is that all redundant parameters are removed. Denote fitted parameters as $(\hat{\beta}, \hat{\sigma}^2)$ such that

$$\hat{\beta} = (X_c' F_c W_c X_c)^{-1}\, X_c' F_c W_c X_c$$

$$\hat{\sigma}^2 = (Y_c - X_c\hat{\beta})' F_c W_c (Y_c - X_c\hat{\beta})/(N_{obs} - p)$$

where $N_{obs} = \sum_{i \in obs(Y)} f_i$ is the number of complete cases, p is the number of parameters, and Y_c, X_c, F_c, W_c are the dependent vector, design matrix and frequency weight, regression weight matrix for complete cases.

The posterior distributions are:

$$\beta | \sigma^2, Y_c, X_c \sim N\left(\hat{\beta}, (X_c' F_c W_c X_c)^{-1}\sigma^2\right)$$

$$\sigma^2 | Y_c, X_c \sim (N_{obs} - p)^{\hat{\sigma}^2 / X^2_{N_{obs}-p}}$$

Let A be the upper triangular matrix of Cholesky decomposition $(X_c' F_c W_c X_c)^{-1} = A'A$

Drawing parameters from the posterior distributions,

- Draw $(\sigma^*)^2$: draw a random value u from $X^2_{N_{obs}-p}$, then $(\sigma^*)^2 = (N_{obs} - p)^{\hat{\sigma}^2/u}$.
- Draw β^*: draw p independent N(0,1) values to create a random vector v, then $\beta^* = \hat{\beta} + \sigma^* A'v$.

then imputing missing values. For i in mis(Y), draw z_i from N(0,1); imputation is $y_i^* = x_i'\beta^* + \frac{\sigma^*}{\sqrt{w_i}} z_i$

Results

Our survey analyses suggest construct reliability and convergence, with the Kaiser-Meyer-Olkin (KMO) measures for all constructs (personality, personal well-being, Facebook usage) ranging from 0.788 to 0.9 (Table 1). In the construct Facebook usage, a Principle Component Analysis (PCA) indicated that two traits, "Do other people present themselves differently in online and offline settings?" (0.391) and "I can be more open online than in real life" (0.487) did not fulfil the KMO criterion of a 0.5 minimum value, and are therefore trimmed from the scale in accordance with (Podsakoff and Organ, 1986). In each PCA analysis, Bartlett's test of sphericity was statistically significant ($p < .0005$), allowing us to reject the null hypotheses. Cronbach's α tests of internal consistency (a standard measure for this type of analysis) showed values ranging from 0.668 - 0.841 (Table 1). It is possible to escalate the unit of analysis if obvious incidents of CMB are identified (Podsakoff and Organ, 1986), as the alternative method of separation of measurement is difficult to implement in crowdwork platforms.

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy		
Personality	Well-being	Facebook usage
0,648	0,900	0,788
Cronbach's α		
Personality	Well-being	Facebook usage
0,603	0,841	0,668

Table 1. Measures of Sampling Adequacy and Internal Consistency

The crowdworkers' results from these surveys indicate replication of (Huppert and So, 2011; John et al, 1991; Ewig, 2011). Workers self-reported current locations in six distinct geographic regions, with the bulk majority of workers reporting locations in North America and India and accordingly the largest language group was English with 285 timelines using predominately English. 73% of workers self-reported to be aged 35 or younger. Gender of the workers is evenly split between women and men, with one non-disclosure and one choice of ‘Other.’ 37% reported being unemployed and 57% completed at least a bachelor’s degree.

Of the 285 English profiles, 282 have profiles with 50 or more words over the lifetime of the profiles. When considering the 285, the average word count per worker is 9379, whereas deleting these three profiles gives an average word count of 11087. This signifies the magnitude of variance in the profiles. Table 2 illustrates some descriptive categories of word counts and usage considering the average and the standard deviation. Again, emoticons and words per profile signify a huge variance. Therefore, the following analysis are normalized for length unless otherwise stated. When text is taken into consideration, only the 282 English profiles with more than 50 words are used.

	μ	Σ
Words per Profile	9379	24367
Emoticons	.05	.07
Unique Words	38	22
+6 Letter Words	16	6

Table 2. Average and Standard Deviation of descriptive aspects of the text

There are some generally interesting results dealing with contact patterns and motivation of use outside of self-representation issues. Workers who use Facebook frequently also update their profiles frequently (r_s(337 = .292, $p < .005$), though those with a higher number of friends have a negative relationship with the frequency of updates (r_s(337 = -.314, $p < .005$). A negative relationship also exists between number of friends and number of updates (r_s(337 = -.252, $p < .005$). A worker with high well-being score has a positive significant relationship with a higher number of Facebook friends (r_s(337 = .112, $p < .041$), but an negative relationship with frequency of updates ($r_s(337) = -.109$, $p < .047$). These results support, yet give a more nuanced understanding to the findings in (Kross et al. 2013) that Facebook usage predicts lowered subjective well-being in young adults.

Family, and on and offline friends on Facebook are a major interest areas for workers. Workers who use Facebook to show what they know and can are less interested in contacting family than all other groups (on and offline friends, unknown people) (Exp(B) = 0.5, p= 0.071). Those who mainly like status updates are most likely to contact family members (Exp(B) = 2.320, p =0.006). Workers who use Facebook in order to be recognized by others and are half as likely to have offline friends on Facebook as the rest of the population (Exp(B) = 0.550, p = 0.085), and are twice as likely to be interested in contacting family members on Facebook (p = 0,067, Exp(B) = 1,989). An exception here is those who want recognition and support from other users: they are half as likely to contact family members (Exp(B) = 0.406, p = 0.011). Men are less interested in maintaining contact with family on Facebook as women (Exp (B) = 0.393, p = 0.001), and those who frequently like videos are twice as likely to use Facebook for contacting their family (Exp (B) = 2.502, p = 0.004). Workers whose profile picture does not show their face are half as likely to want to contact offline friends and are more interested in finding unknown online friends (Exp(B) = 0.413, p = 0.007), as well as workers who agree with the statement "I can determine myself what I do or do not show others" (Exp(b) = 1.344, p= 0.033).

Identifying Self-Representation

By concentrating on a selection of sentiment categories known to correlate with deception, personality, and confidence, we construct and then cluster individual's propensity to self-represent in their online social media persona (Buckels et al., 2014; Tausczik and Pennebaker, 2010; Newman, 2003; Yarkoni, 2010). We recognize that these indicators are unlikely to be the only psychometrics that could be indicative of self-representation, but they are the most thoroughly researched and thus the most appropriate for this analysis. Deceptive profiles as identified in (Newman et al. 2003) were assessed by first establishing the mean of the LIWC categories first person singular, motion, exclusion, and negative emotion. Two cut-offs were employed, by adding the first and second standard deviation to the average.

Those who employ above average negative emotion, more motion words, fewer exclusion words, and less first-person singular are considered to display potential signs of lying. Fitting this description are 96 worker profiles. We further separate these profiles in order to use them as a control factor in the data.

In order to verify that the data is reliably reported the hypotheses H1 and H4 on the relationships between personality and happiness are established (Table 3). If confirmed, H 1 and 4 the data can be considered to reflect population attributes from social psychology literature (Diener & Suh 1997; Hall, Caton, et al. 2013). If confirmed, our assumptions are that Hypotheses 2, 3, 5, and 6 should also be confirmed; otherwise, issues of self-representation in the data are likely evident in the data.

		1 or 2 Tailed?
H1	There is a positive relationship between human flourishing and extraversion	1
H2	There is a relationship between extraversion and positive writing traits	2
H3	There is a relationship between human flourishing and positive writing traits	2
H4	There is a negative relationship between human flourishing and neuroticism	1
H5	There is a relationship between neuroticism and negative writing traits	2
H6	There is a relationship between human flourishing and negative writing traits	2

Table 3. Hypotheses listed for testing reliability of the data

For the one-tailed hypotheses of a positive relationship existing between human flourishing and extraversion and a negative relationship existing between neuroticism and human flourishing, both hypotheses are strongly confirmed ([$r_s(282) = .357$ $p < .0005$] [$r_s(282) = -.263$ $p < .0005$]). Hypotheses 2 and 5 were further broken down from “writing traits” into their respective LIWC categories (Table 4). Here only H5b can be confirmed, namely there is a relationship between neurotic personality types and expressed anxiety from posts on Facebook. This indicates that self-representation is likely to be higher with those who self-identify as extraverts, whereas neurotic personality types do leave some digital indicators of their personality, whilst still maintaining some self-representation aspects.

		ρ	p	✓/≈/¬
H2	There is a relationship between extraversion and positive writing traits	-	-	¬
H2a	There is a relationship between extraversion and (written) positive emotion	-.019	.751	¬
H2b	There is a relationship between extraversion and positive feelings	-.031	.598	¬
H2c	There is a relationship between extraversion and optimism	-.016	.795	¬
H5	There is a relationship between neuroticism and negative writing traits	-	-	≈
H5a	There is a relationship between neuroticism and (written) negative emotion	.069	.402	¬
H5b	There is a relationship between neuroticism and anxiety	.120*	.043	✓
H5c	There is a relationship between neuroticism and anger	.061	.307	¬
H5d	There is a relationship between neuroticism and sadness	.050	.398	¬

Table 4. Hypotheses on the relationships between personality and LIWC categories

Hypothesis 3 and 6 were likewise expanded to correspond human flourishing to related LIWC categories (Table 5). Here there is likewise one significant relationship, that of human flourishing and optimism. Here the case is that extraverts will accurately portray their propensity to feel optimistic in their writing, though nothing else, where neurotics self-represent their traditionally negative views out of their Facebook information. As hypotheses 1 and 4 are confirmed, whereas only H2b and H3c are confirmed of the remaining 18, we can state with some confidence that workers' have (either on purpose or inadvertently) systematically self-represented themselves on their Facebook profiles. A further note is that when statistically controlling for lying, the weak significance of H2b and H3c disappears. This could be a confirmation that lying is not in fact the same as self-representation.

		ρ	p	✓/≈/¬
H3	There is a relationship between human flourishing and positive writing traits	-	-	≈
H3a	There is a relationship between human flourishing and (written) positive emotion	.102	.088	¬
H3b	There is a relationship between human flourishing and positive feelings	.030	.612	¬
H3c	There is a relationship between human flourishing and optimism	.144*	.015	✓
H6	There is a relationship between human flourishing and negative writing traits	-	-	¬
H6a	There is a relationship between human flourishing and (written) negative emotion	.016	.785	¬
H6b	There is a relationship between human flourishing and anxiety	-.035	.557	¬
H6c	There is a relationship between human flourishing and anger	.029	.625	¬
H6d	There is a relationship between human flourishing and sadness	-.025	.682	¬

Table 5. Summary: Hypotheses on the relationships between happiness and LIWC categories

Sentiment analysis shows that the relative frequency of positive and negative emotion shifts across the lifespan of a timeline, allowing for pattern establishment and estimation of transient mood states. Workers generally communicate their positive emotions more frequently (an average of 4.25% of all text), where negative emotions in Facebook are hardly communicated (1.2% of all data), regardless of personality type and in line with the results of (Qiu et al. 2012). As 60% more words within the LIWC dictionary are associated with negative sentiment the social posturing aspects here are clear. As the scales utilized had minimal social desirability and are balanced in positive and negative words (see Appendix I) in line with (Podsakoff et al. 2012; Qiu et al. 2012), this study can identify “displays of positive emotion” and “hiding negative emotion” as forms of self-representation bias.

The analysis also looked at expressed confidence as a measure of self-representation. This is measured by the frequency in usage of first person singular and third person plural; where people that are more confident use “I” words less than “We” words. Here we tested the demographic groups established in the survey with an ANOVA (Figure 2) and found a significant difference in gender (Gender $F(2,279) = 11.893$, $p < .0005$; Wilks' $\Lambda = .921$; partial $\eta^2 = .079$). Our findings cannot reject a difference between third person plural between men and women (First Person Plural (We) $F(1,280) = .643$, $p = .423$; partial $\eta^2 = .002$), whereas first person singular has a significant difference in gendered usage (First Person Singular (I) $F(1,280) = 23.405$, $p < .0005$; partial $\eta^2 = .077$). There was homogeneity of variance-covariance matrices, as assessed by Box's test of equality of covariance matrices ($p = .002$). This supports emerging findings[2] that women express less confidence than men do, and thereby does not support overt self-representation specific to online social networks.

[2] http://www.theatlantic.com/features/archive/2014/04/the-confidence-gap/359815/

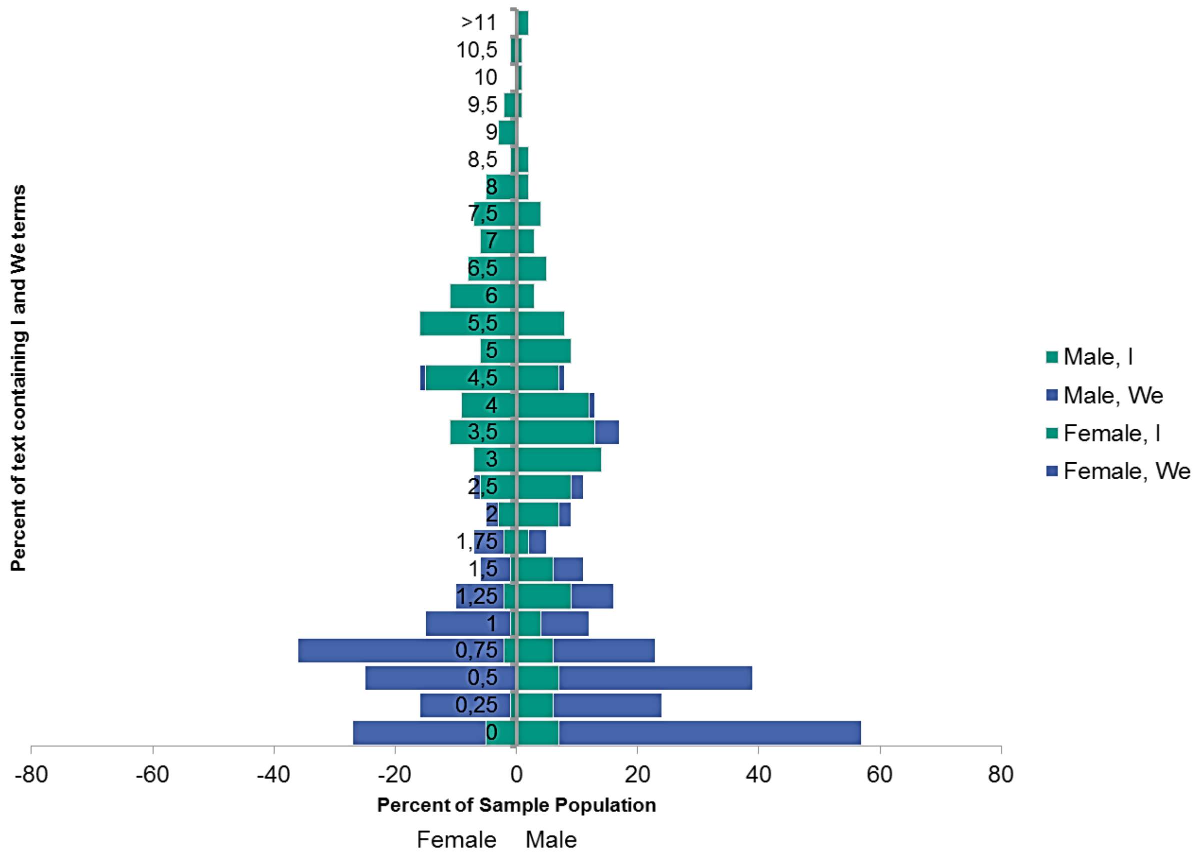

Figure 2. Gendered usage of confident statements on Facebook profiles

Personality as a tool for mitigating self-representation

Regressing the 136 variables[3] of survey responses and sentiment categories on each of the five personality traits of the Five Factor model (John et al. 1991) creates meritorious model fits (Table 6; component tables are found in Appendix 1), without overt signs of overfitting. The multivariate models are statistically significant for each trait, with some overlap of the variables predicting the Five Factor traits. Considering sizeable correlations between predictor groups, the unique variance explained by each of the variables indexed by the squared semipartial correlations is low. In no case was there an instance of Cook's D being larger than one, so all outliers were handled within the data rather than trimmed. The coming section is a short discussion of the predictors of each trait, with predictors grouped

[3] We excluded punctuation and the corresponding four Five Factor traits per instance from this analysis.

by measurement instrument then listed by weight. In order to constrain the number of variables, only the first ten items significant at the ($p < .001$) level are reported.

Trait Name	**Reference Model**	**Ensemble**	**Δ**
Openness	78.5	77.3	1.2
Conscientiousness	69.4	64.3	5.1
Extraversion	77.8	69.5	8.3
Agreeableness	71.4	71.0	0.4
Neuroticism	75.9	68.9	7.0
Average	*74.6*	*70.2*	*4.4*

Table 6. Prediction accuracy per model on Five Factor Personality traits, boosted (10 component models) using best-subsets

Openness has the high prediction accuracy at 78.5%, and is a very stable prediction given the low difference indicates that the prediction is relatively stable. Highly significant are the survey categories meaning, self-esteem, engagement, competence, optimism, positive emotion, and resilience; the country of origin of the worker; and the sentiment category “feelings.”

With the lowest prediction accuracy (69.4%) and a medium model difference (5.1%), conscientiousness must be considered less reliable. The sentiment categories, ‘friends’, ‘down’, and ‘fillers’; survey responses ‘pictures that do not show me’, number of friends, ‘I understand quickly how others perceive me’, assent to ‘People should present themselves on online social networks as the same person as they are offline’, and using Facebook to give and get information, and the survey measurement resilience and positive relationships are the most relevant predictors.

Extraversion (77.8% accuracy and the largest difference of 8.3%) is related to the survey items competence, self-esteem, meaning, optimism, positive emotion, vitality, and resilience; country of origin;

and the survey responses ‘I understand quickly how others perceive me’ and managing Facebook profiles with displays of albums.

Agreeableness has the lowest deviation (0.4%) and an accuracy of 71.4%, indicating high reliability. Highly significant are the survey items resilience, meaning, self-esteem, and competence; country of origin; the sentiment categories ‘friends’, ‘inhibition’, ‘feelings’, and ‘assent’; and declination of ‘I can be who or what I want on my Profile page.’

The final trait, Neuroticism, also has a high deviation between models (7%), but a good performance (75.9 % accuracy). The most significant survey items are resilience, self-esteem, emotional stability, vitality, and optimism; the survey responses ‘I use Facebook to spy on others’, managing presentation of self with pictures that are not of them, using Facebook to observe other people, and liking videos on Facebook. Additionally, the sentiment category ‘feelings’ is highly significant.

A sub analysis concentrating on only sentiment categories also shows that topical discussions have high prediction value for the Five Factor model (Figure 3). Highly significant for openness are the sentiment categories ‘sports’, ‘religion’, ‘feelings’, ‘music’, ‘fillers’, and ‘TV’, where sports, music, fillers and TV have a positive association with openness, feelings have a negative association, and religion has a U-shaped relationship very low and high openness scores have a positive association, but mid-scores have a negative association. Conscientiousness displays that ‘religion’, ‘friends’, ‘TV’, ‘inhibition’, and ‘music’ are positively related, and ‘fillers’ is negatively related. Extraversion is positively related to ‘inhibition’ and ‘TV’, and negatively related to ‘friends, ‘sports’, and ‘down’. Agreeableness’ highly significant sentiment categories are negative relationships with ‘inhibition’ and ‘death’, and a positive relationship with ‘friends’. The final trait, neuroticism finds ‘religion’, ‘friends’, ‘TV’, ‘inhibition’, and ‘music’ being positively related and ‘fillers’ and the survey response resilience being negatively related.

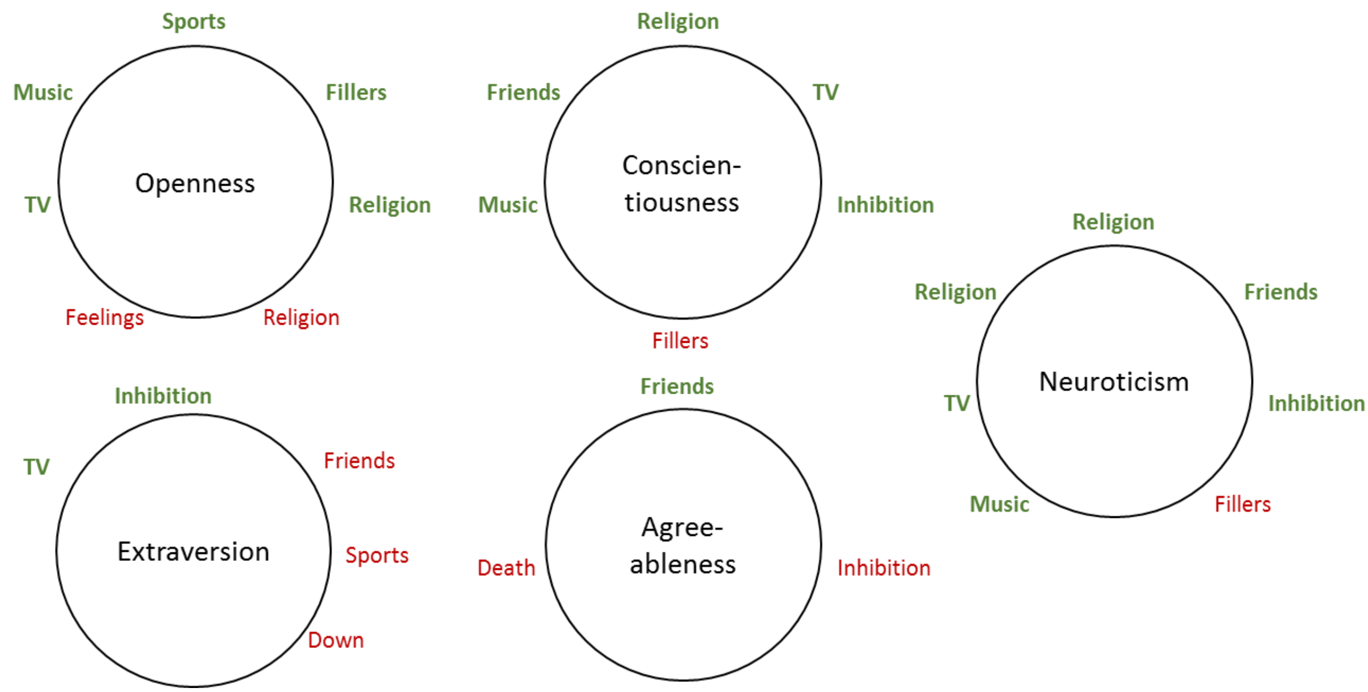

Figure 3. Five Factor Model mapped to positive (Bolded green) and negative (red) relationships of LIWC sentiment categories with high predictor strength

Several sentiment categories dominate the results; while surprising at first glace, when the medium of data is considered, the findings are less surprising. Facebook is a medium to exchange news and ideas, and while more reflective in nature and practice than Twitter (Dodds et al. 2011), is still essentially used as a short information service to connect people (Wilson et al. 2012; Hampton et al. 2011). Inhibition is also very common, and suggests that workers (consciously or not) are in fact utilizing vocabulary of inhibition on their Facebook profiles. This could be indicative of active-self representation.

Discussion

This research finds that self-representation in online social media is an identifiable phenomenon and must be addressed in the collection and analysis of such data. The surveys in use are empirically validated and the data collection and processing found that 82% of the sample did not violate constraints suggested in (Podsakoff et al. 2003; Podsakoff et al. 2012). The text samples were generated in a way which did not induce Type I errors (González-Bailón et al. 2014), and Type II errors are not in scope for this type of work. Whilst lying was identified in the text-based sample, the control measures noted above mitigated this behavior in the survey items. Lying profiles were isolated, and used as a control item.

Self-representation was identified in a number of indicators. While the survey-only results show a replication of literature, the survey to text results cannot replicate the findings that extraversion is a predictor of well-being, and neuroticism has a negative relationship with well-being. This is true of both 'honest' and 'dishonest' Facebook profiles. Positive affectivity and withdrawn negative emotions are identifiable patterns across all workers' profiles. However, confidence can be identified and follows expected patterns across genders. Male participants appear more confident in their written profiles than females. As this is a finding in emergent literature, this cannot be understood as an overt measure of self-representation.

Given the highly clustered, trivial nature of the sentiment-based predictors, a tempting statement is that the data is not appropriate for the task. However, discernable patterns are present. Moreover, given the platform, the topics discussed are a reasonable (albeit, surprising) output. The topical basis of the other predictors conceptual themes of workers' discussions, and neatly create psychological profiles that link online and offline personality. Especially the strength of inhibition in four of five of the Five Factor model suggests that the participants have been displaying reticence in their Facebook profiles. In future research, when the linguistic habits are identified, the researcher is able to control for this in data preparation or as a control factor in the calculation, e.g., as a dummy variable in regression models.

Conclusion, Limitations, Future Work

The stated aims of this research are twofold, both establishing the relationship between offline and online personalities, and mitigating of biases in surveys and in publically sourced data. In accomplishing these goals, this research creates a generally applicable, bias-free method for the design of cross-disciplinary methods, and the analysis of social media data. In a systematic manner, this research detailed the experimental design, data collection, and analysis. Common method biases are addressed at each step of the process and appropriately eliminated when isolated. The method allows for replication by careful detailing of the steps and processing of data. A strength of this study is its careful review of the findings from recent so-called "cyber psychology" literature. The replications found from this basis show that the methods of eliciting survey data from crowdworkers is reliable, and indicates that the sourced data is robust when compared to the other authors' findings. A major contribution is addressing method biases in the harvesting and analysis of social media data. This research utilizes the entire data stream per profile, mitigating Type I bias. With personality and mood validated and a sentiment analysis performed on the lifespan of a user's Facebook timeline, we can now measure the propensity of a user to portray themselves in opposition to their truthful, psychological baseline. It also names common markers of the phenomena of self-representation based on simple sentiment categories and psychometrics that allow researchers to mitigate its effects in future research.

This work is not without fault. A major limitation is the sample size, which disallows larger statements about subgroups as the non-English samples are too small for meaningful statistics. Another drawback is that the results are tailored to Facebook – the findings of this study are unlikely to generalize to professional networking, microblogs, or visual media sites. A known issue of Natural Language Processing is that the state of the art tools are unable to capably handle sarcasm and irony (Tsur & Rappoport 2010), which has unknown effects across the lifespan of a Facebook timeline. A final drawback of the tools utilized is that LIWC is unable to be called from the command line, which limits the researcher in ability to look at highly granular levels like post by post. A concluding remark on limitations is related to privacy. While the study obtained informed consent of its workers, the open question remains if workers truly understood the amount of information that was being given in the task.

Natural extensions of this research are closely linked to its limitations. More granular review of the data is possible, given that a version of the text analysis tool is extended in a way that easily allows post-by-post analysis. Cross-platform analysis of the same user for their various public profiles would give future work a more nuanced view in the ways that social media users self-represent in difference audiences. Such a work would fill research gaps in 'best' platform usage for information disbursement, creation, and influence, as well as impact for a given network. That in mind, a network analysis of users, and textured understanding of how users cluster and complement within a network would be a good area of future research.

We propose that policy-makers, researchers as well as community managers can apply this method to their analyses of publically sourced data in order to mitigate the effects of various phenomena, including trolling, social desirability, and acquiescent behaviours (e.g., the spiral of silence). Such an approach has diverse applications in policy and community management, in that it allows for a new, accurate measurement system from which to deduce from publically accessible text onto the general population. With self-representation identified and removed, policy-makers and researchers guaranteeing the validity of their measurements of people without necessitating expensive survey methods.

References

1. Auer, M.R. (2011). The Policy Sciences of Social Media. *The Policy Studies Journal*, 39(4), pp.709–736.
2. Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), 351-368.
3. Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*. http://dx.doi.org/10.1016/j.paid.2014.01.016
4. Chung, C.K. & Pennebaker, J.W., 2014. Counting little words in Big Data: The Psychology of Communities, Culture, and History. In J. Forgas, O. Vincze, & J. Laszlo, eds. *Social Cognition and Communication*. New York, New York, USA: Psychology Press, pp. 25–42.
5. Das, S., & Kramer, A. (2013). Self-censorship on Facebook. *Proc. of ICWSM 2013*, 120-127.

6. Diener, E. & Suh, E., 1997. Measuring quality of life: economic, social, and subjective indicators. *Social Indicators Research*, 40, pp.189–216.
7. Dodds, P.S. et al., 2011. Temporal Patterns of Happiness and Information in a Global Social Network : Hedonometrics and Twitter. *PloS one*, 6(12).
8. Ellison, N., Heino, R. & Gibbs, J., 2006. Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. *Journal of Computer-Mediated Communication*, 11(2), pp.415–441.
9. Ewig, C., (2011) "Identität und Soziale Netzwerke – StudiVZ und Facebook" In Social Media: Theorie und Praxis digitaler Sozialität. Caja Thimm and Mario Anastasiadis (eds) Berlin: Peter Lang Internationaler Verlag der Wissenschaft.
10. Fowler, R.L., 1987. Power and Robustness in Product-Moment Correlation. *Applied Psychological Measurement*, 11(4), pp.419–428.
11. Goffman, E., 1959. The Presentation of Self in Everyday Life. , pp.135–146.
12. González-Bailón, S., Wang, N., Rivero, A., Borge-Holthoefer, J., & Moreno, Y. (2014). Assessing the bias in samples of large online networks. *Social Networks*, 38, 16-27.
13. Hall, M., Caton, S., & Weinhardt, C. (2013). Well-being's Predictive Value. In A. A. Ozok & P. Zaphiris (Eds.), *Proceedings of the 15th International Conference on Human-Computer Interaction (HCII)* (pp. 13–22). Berlin: LNCS, Springer Verlag.
14. Hall, M., Glanz, S., Caton, S., & Weinhardt, C. (2013). Measuring Your Best You: A Gamification Framework for Well-being Measurement. In *Third International Conference on Social Computing and its Applications* (pp. 277–282). Karlsruhe, Germany: IEEE. doi:doi:10.1109/CGC.2013.51
15. Hampton, K.N. et al., 2011. *Social networking sites and our lives*, Pew Research Center, Washington, DC.
16. Hampton, K.N. et al., 2014. *Social Media and the Spiral of Silence*, Pew Research Center, Washington, DC.
17. Hoever, A. (2010). Strategie und Konzepte der Selbstdarstellung auf Social Network Services am Bespiel Facebook. *Berlin Qualitative Methods Symposium*. http://www.qualitative-forschung.de/methodentreffen/archiv/poster/poster_2010/hoever.pdf
18. Huppert, F.A. & So, T.T.C. (2011). Flourishing Across Europe: Application of a New Conceptual Framework for Defining Well-Being. *Social Indicators Research* 110(3), 837–861.
19. IBM_a, 2011. *IBM SPSS Advanced Statistics 22*.
20. IBM_b, 2011. *IBM SPSS Regression 22*.
21. John, O.P., Donahue, E.M., & Kentle, R.L. (1991). The big five inventory—versions 4a and 54. Berkeley, CA.
22. Kramer, A., 2010. An Unobtrusive Behavioral Model of "Gross National Happiness." In *Proceedings of the 28th International Conference on Human factors in Computing Systems*. ACM CHI pp. 287–290.
23. Kramer, A., 2012. The spread of emotion via Facebook. In *Proceedings of the 2012 ACM Annual conference on Human Factors in Computing Systems - CHI '12*. New York, New York, USA: ACM Press, pp. 767–770. Available at: http://dl.acm.org/citation.cfm?doid=2207676.2207787.
24. Kramer, A., Guillory, J.E. & Hancock, J.T., 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1320040111 [Accessed June 3, 2014].
25. Kross, E. et al., 2013. Facebook use predicts declines in subjective well-being in young adults. *PloS one*, 8(8), p.e69841. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3743827&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

26. Kushin, M.J. & Yamamoto, M. (2010). Did Social Media Really Matter? College Students' Use of Online Media and Political Decision Making in the 2008 Election. *Mass Communication and Society*, 13(5), pp.608–630.
27. Linville, P. W. (1985). Self-complexity and affective extremity: Don't put all of your eggs in one cognitive basket. *Social Cognition*, 3(1), 94-120.
28. Qiu, L., Lin, H., Leung, A.K., & Tov, W. (2012). Putting their best foot forward: emotional disclosure on Facebook. *Cyberpsychology, Behavior and Social Networking*, 15(10), 569–72.
29. Hogan, B., 2010. The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, 30(6), pp.377–386.
30. Mehra, A., Kilduff, M. & Brass, D.J., 2001. The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly*, 46(1), pp.121–146.
31. Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, (29), 665-75.
32. Noelle-Neumann, E., 1974. The Spiral of Silence: A theory of public opinion. *Journal of Communications*, 24(2), pp.43–51.
33. Pennebaker, J.W. et al., 2007. *The Development and Psychometric Properties of LIWC2007*, Austin, TX: University of Texas, Austin.
34. Pennebaker, J.W., Mehl, M.R. & Niederhoffer, K.G., 2003. Psychological aspects of natural language use: our words, our selves. *Annual review of psychology*, 54, pp.547–77.
35. Podsakoff, P. M., Mackenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research : A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), 879–903. doi:10.1037/0021-9010.88.5.879
36. Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539–69. doi:10.1146/annurev-psych-120710-100452
37. Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531-544.
38. Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, 7(3), 321-326.
39. Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology, 29(1), 24-54.
40. Thelwall, M. et al., 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), pp.2544–2558.
41. Tsur, O. & Rappoport, A., 2010. ICWSM – A Great Catchy Name: Semi-Supervised Recognition of Sarcastic Sentences in Online Product Reviews. *ICWSM*, pp.1–9.
42. Utz, S., Tanis, M. & Vermeulen, I., 2012. It is all about being popular: the effects of need for popularity on social network site use. *Cyberpsychology, behavior and social networking*, 15(1), pp.37–42.
43. Wang, N. et al., 2014. Can Well-Being be Measured Using Facebook Status Updates? Validation of Facebook's Gross National. *Social Indicators Research*, 115(1), pp.483–491.
44. Wilson, R.E., Gosling, S.D. & Graham, L.T., 2012. A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science*, 7(3), pp.203–220.
45. Yang, H., 2013. The Case for Being Automatic: Introducing the Automatic Linear Modeling (LINEAR) Procedure in SPSS Statistics. *Multiple Linear Regression Viewpoints*, 39(2), pp.27–37.
46. Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, 44(3), 363-373.
47. Zhao, S., Grasmuck, S., & Martin, J. (2008). Identity construction on Facebook: Digital empowerment in anchored relationships. *Computers in Human Behavior*, 24(5), 1816–1836.