Using Twitter for Demographic and Social Science Research: Tools for Data Collection[1]

Tyler H. McCormick
Hedwig Lee
Nina Cesare
Ali Shojaie

Working Paper no. 127
Center for Statistics and the Social Sciences
University of Washington

April 8, 2013

[1] Tyler H. McCormick is Assistant Professor of Statistics and Sociology, Center For Statistics and the Social Sciences, University of Washington, Box 354320, Email: tylermc@uw.edu. Hedwig Lee is Assistant Professor of Sociology, Department of Sociology, University of Washington, Box 353340, Email: hedylee@uw.edu. Nina Cesare is a graduate research assistant, Department of Sociology, University of Washington, Box: 353340, Email: ninac2@uw.edu. Ali Shojaie is Assistant Professor of Biostatistics, Department of Biostatistics, University of Washington, Box 357232, Email ashojaie@u.washington.edu.
We would like to acknowledge the support of the Army Research Office, a UW Center for Statistics and the Social Sciences Seed Grant, and a UW School of Public Health Emerging Challenges Seed Grant. We also appreciate technology support from the Center for the Studies in Demography and Ecology at the University of Washington.

Abstract

Despite recent interest in using Twitter to examine human behavior and attitudes, little work has been done to develop systematic ways of collecting Twitter data for social science research. Further, gleaning key demographic information about Twitter users, a key component of much social science research, remains a challenge. This paper develops a scalable, sustainable toolkit for social science researchers interested in using Twitter data to examine behaviors and attitudes, as well as the demographic characteristics of the populations expressing or engaging in them. We begin by describing how to collect Twitter data on a particular population – in this case, individuals who do not plan to vote in the 2012 U.S. presidential election. We then describe and evaluate a method for processing data to retrieve demographic information reported by users that is not encoded as text (e.g., details of images) and assess the reliability of these techniques. We end by assessing the challenges of this data collection strategy and discussing how large-scale social media data may benefit demographic researchers.
KEY WORDS: Twitter; social media; big data, demography; data collection; crowdsourcing; Amazon Mechanical Turk

1. INTRODUCTION

Twitter and the rise of "Big Data"

Social media data, such as Twitter and Facebook, provide exciting opportunities that, according to a recent issue of the American Sociological Association (ASA) magazine, can "open up a new era" of social science research (Golder and Macy 2012). These new communication platforms afford the opportunity to examine social data on a variety of topics on a massive scale and to collect these data over very short periods of time. Social media websites and the data extracted from them have gained a growing interest among many researchers who are attempting to shows how these platforms influence or reflect social relationships and behavior (Brickman Bhutta 2012; Golder and Macy 2011; Heaivilin, Gerbert, Page, and Gibbs 2011; Lowe, Barnes, Teo, and Sutherns 2012; Moreno, Grant, Kacvinsky, Egan, and Fleming 2012; Valkenburg, Peter, and Schouten 2006). Though a few social science researchers have begun to use Twitter to document changing moods and other sentiments and opinions on the aggregate level (Diakopoulos and Shamma 2010; Golder and Macy 2011; Naaman, Becker, and Gravano 2011; Reips and Garaizar 2011; Yardi and Boyd 2010), the potential of such data for demographic research has yet to be realized.

Social media data represent a new data collection paradigm for social science research. These data share come features with more well-researched data collection mechanisms, such as surveys or structured observations, but also contain new features. Surveys, for example, ask respondents to recall behaviors or sentiments retrospectively, whereas social media data afford the opportunity to observe behaviors and human interaction in real-time and on a large scale. With appropriate infrastructure, scientists can analyze and begin presenting results within a matter of months (or sooner), rather than the years typically required for a survey. Social media data are also share some characteristics of observational or ethnographic work. Specifically, social media data allow researchers to collect reports of behaviors that are unsolicited and unprompted by a researcher. One could even argue that these data provide a better reflection of day-to-day social experiences. Indeed, Twitter interactions have been described as persons "want[ing] to know what the people around them are thinking and doing and feeling, even when co-presence isn't viable" and "shar[ing] their state of mind and status so that others who care about them feel connected (boyd 2009)." Unlike previous observational work, however, the context of interactions on social media data can be captured and stored. Once an in-person interaction has passed, for example, it cannot be reconstructed and a researcher doing observational work is left with only her or his perceptions (and notes) on the social context. The social context of an interaction on social media data, however, is preserved and can be reviewed multiple times and passed to other interested researchers.

Despite these possibilities, social scientists often see such data as inaccessible for social science research and solely relevant to computer and physical scientists. The same ASA article (Golder and Macy 2012: 7) laments, "most of the social and behavioral science using online data is coming from computer and information scientists who do not always have the training required to ask the right questions, or to recognize unfounded assumptions and socially unjust ramifications."

A further hindrance arises as currently each investigator must devise a unique sampling strategy for social media data collection. As of now, very little social science research has been able to systematically collect data from Twitter (Heaivilin, Gerbert, Page, and Gibbs 2011; Krishnamurthy, Gill, and Arlitt 2008; Naaman, Becker, and Gravano 2011). This prospect is especially challenging given the numerous differences between Twitter data and the surveys collected using traditional sampling techniques. Using traditional surveys, for example, researchers see comparatively few respondents but have a great deal of control over what information respondents provide. Under these conditions, respondents provide information of interest to the researchers, but the limited sample size may not produce enough variability to study less commonly observed phenomena in their entirety (e.g., self-reports of suicide attempts, eating disorders, or HIV positive status). Data from Twitter in contrast, is completely unelected but offers unprecedented exposure to variability. On the other hand, the uncontrolled nature of information sharing on Twitter necessitates that such data be verified.

In addition to challenges associated with sampling, it is difficult to gather demographic data from text based blogs and microblogs such as Twitter. Demographic information is at the heart of most social science analysis. It is often important that researchers are able to utilize information on race, age and gender to examine patterns in attitudes and behaviors. This is also a challenge for Twitter data, where individuals are not asked to respond to questions about demographic. Removing the actual and perceived barriers that prevent social scientists from using social media data offers new research opportunities for social scientists and increases the potential for interdisciplinary research between computer scientists or statisticians with social and behavioral scientists, thus increasing the potential of studying complex social problems.

This paper will describe the process of developing a scalable, sustainable infrastructure that facilitates access to demographic information from Twitter data. Furthermore, it seeks to encourage social scientists to consider Twitter as a valuable source of demographic and other behavioral/social information to answer relevant social science questions. To help illustrate this data collection process, we examine a specific behavior, reporting the intention to not vote in the 2012 presidential election.

In the following sections, we outline our three-pronged approach of data extraction, processing, and analysis. We begin with an introduction to Twitter and a discussion of the resources and challenges, associated with using data extracted from Twitter. Next, since Twitter users typically do not report demographic information directly, we describe a processing strategy that allows us to gather this information from users' profile photos. At the heart of the strategy is a framework for using Amazon Mechanical Turks[i] to efficiently code large volumes of images. We conclude by addressing the benefits and challenges associated with this method of data collection, as well as the potential for future research using demographic data obtained from Twitter.

Applications of Twitter data

General applications

Twitter provides an inexpensive and convenient source of data about users' opinions, interactions, and reported behaviors. It may be utilized, for example, by researchers who seek to examine large-scale processes of contagion, track preferences and/or opinions among broad audiences, examine behaviors and attitudes where social desirability bias in an official survey may occur (e.g., racist attitudes, voting behavior or anti-immigrant sentiments) (Belli et al. 2009; Holbrook and Krosnick 2010; Janus 2010; Tourangeau and Yan 2007), analyze collective experiences based on a timely event (e.g., teacher strikes, terrorist attacks or natural disaster), gather large amounts of data on hard-to-reach populations, and pretest to see if attitudes and behaviors not present in current surveys are evident among particular population subgroups.

In addition to attitude and trend tracking, Twitter data can also prove useful in the field of population health. Achrekar and colleagues (2011), for example, track Twitter posts containing mentions of influenza in order to create a real time illustration of the spread of the illness. Heaivilin and colleagues (2011) use Twitter data as a means of gathering information on the prevalence of oral health problems and the actions taken to remedy them. Golder and Macy (2011) approach Twitter from a mental health perspective and use data collected from this platform as a means of tracking how sleep patterns and day length impact individuals' moods. These authors note that the candidness of Twitter users in discussing personal matters – such as oral health or emotional status – suggests that healthcare providers may begin to use this platform as a tool for monitoring the public's health and communicating with patients.

Regardless of its specific application, these studies and others suggest that Twitter provides a cost effective means of developing a broad understanding of a populations' activities and attitudes. In other words, the content of users' Tweets provides insight into what Naaman, Becker and Gravano (2011) call "social awareness streams." This source of organically created and automatically archived human data allows researchers to see what people are doing, what they are saying, and how they feel about particular issues as these actions and thoughts arise. This is an unprecedented form of data for social scientists with broad research potential but marked challenges due to its relative unfamiliarity. The ability to systematically gather demographic data from this source would greatly expand the potential for research that seeks to capitalize on the availability of Twitter data.

Using Twitter for Political Analysis

One popular application of Twitter data is political analysis and election forecasting. Prompted by promising analyses of political opinion trends within the blogosphere and other social media outlets, some scholars have explored whether Twitter – despite its limited content – provides a useful outlet for examining political preferences as they develop. There are a number of interesting applications of Twitter data in this burgeoning body of research. Tumasjan and colleagues (2010), for example, collected 100,000 Twitter messages prior to the 2009 German federal election, analyzed these tweets for mentions of party affiliation and positive or negative sentiment, and were able to effectively conclude that the opinion trends reflected in their data paralleled the results of the election. Politicians themselves have noted the popularity of Twitter as a tool for political exchange and many now use this platform as a means of reaching out to potential constituents, though the precise effects of doing so on a particular candidate's electoral success remain inconclusive (Lassen and Brown 2011). In addition to predicting electoral outcomes, Twitter also provides a promising tool for examining the opinion landscape of a nation in regards to political issues. For instance, coding the political content and sentiment of tweets related to a particular issue and tracking the responses and sharing patterns for these tweets allows researchers to illustrate the presence and growth of polarized spheres on Twitter. Conover and colleagues (2011), for example, accomplish this task by harvesting pockets of political discourse on Twitter, coding each tweet for left or right wing sentiment, and mapping interactions between these opposing groups.

Although many researchers have found ways to examine political trends using Twitter data, these studies unanimously lack thorough consideration of the demographic trends underlying them. The addition of this dimension could greatly benefit researchers looking to predict political outcomes or track changing tides in political opinion or participation among individuals of particular groups. This trend is of course not unique to political analysis; the same can be said of other research that uses Twitter data for social science research. The addition of demographic data to projects that draw upon social media data for social science research could greatly expand the explanatory and predictive power of these analyses. The following sections propose a method for gathering demographic data in conjunction with information on collective voting habits as a means of expanding upon existing applications of social media data within social science research.

2. CASE STUDY: USING TWITTER TO LOOK AT NON-VOTERS

The objective of this study is to establish a systematic means of gathering demographic information from Twitter users, thus overcoming an important limitation of this rich data source. The first step in this process involves choosing a focal population. This analysis will focus specifically on extracting the demographic characteristics of individuals who express a refusal to vote in the 2012 presidential election. The subsequent steps involve establishing systematic means of extracting data from Twitter, utilizing elements of the Twitter profiles included in the data crawl, organizing and cleaning the data, and extracting demographic information about the users who post these anti-voting tweets. Following these steps, the reliability of the demographic information will be assessed. Finally, we will provide basic descriptive statistics of individuals on Twitter who reported that they would not vote (in the 2012 US presidential election) and compare them to national estimates of people who report that they did not vote (Pew 2012).

This analysis stands out from similar analyses in two important ways. First, this analysis examines not only *what* is said, but also *to whom* these opinions belong (i.e., the demographic characteristics of Twitter user who report not voting). This methods presented here utilize the personal content contained within individuals' Twitter profiles and develop a way to systematically add a layer of previously unavailable demographic information. Doing so will allow social scientists to not only track trends and opinions using Twitter, but to examine the demographic characteristics of opinion-based networks and predict behaviors and attitudes based on social connections as well. Second, the nature of Twitter reveals novel features of social dynamics not captured with other platforms. Unlike social networking platforms, such as Facebook, which rely on mutual connections (two individuals cannot be connected unless both parties confirm), Twitter users' ties are not always reciprocal and not always forged around existing connections. Furthermore, Twitter users have the capacity to control the context of their presentation by concealing their real names and making their profiles unsearchable (Hogan 2010). Therefore, the self-presentation techniques of Twitter users are distinct from those of other social networking platform users in that Twitter users have a tendency to partially disregard their audience and tweet with a level of disclosure and authenticity not present on other social media websites (Marwick and boyd 2010).

Properties of Twitter

Twitter is a microblogging platform that allows users to record their thoughts in 140 characters or less. The text-based content of these messages may include personal updates, humor, or thoughts on media and politics. This concise format allows users to update their blogs multiple times per day, rather than every few days, as is the case with traditional blogging platforms (Java, Song, Finin, and Tseng 2007). Besides projecting their thoughts independently, users can communicate with one another either through private messages, by re-tweeting one another's tweets, or by using the *@reply* command. They may also contribute to broader conversations by including a *hashtag* identifier in their tweet. Tweets from those whom the user follows are displayed as a sequential feed that is updated in real time. Twitter was originally intended to be used via mobile devices (specifically via text message) to facilitate frequent updating, but tweets can also be sent using other internet capable devices, including smart phones, tablets and computers. This mobile interface helps ensure that Twitter users post only short messages and have the capacity to update multiple times per day.

Self-presentation (Goffman 1959) on Twitter is developed through active conversation as well as the maintenance of personal profiles. To generate this conversation, Twitter users project their thoughts toward an imagined audience of networked individuals (Marwick and boyd 2010), some of whom bear reciprocal ties to the users themselves and some of whom do not. This is different than other social networking sites such as Facebook, in which all users are reciprocally tied to one another and disclose information mutually. This interesting mix of public and private attention requires users to maintain a balance between transparency and authenticity in the material they choose to tweet (Marwick and boyd 2010). It is important to note, however, that such considerations of disclosure do not apply to the entire Twitter population. According to the social media analytics platform Beevolve, only 11.8% of all Twitter users choose to "protect" their accounts – meaning the tweets associated with these accounts are only viewable by approved followers. Nonetheless, the strong majority of Twitter users (88.2%) manage a public presence.

There are some who debate whether Twitter provides actual insight into collective experiences or whether the majority of Twitter content is "pointless babble" with no real substantive meaning (boyd 2009). While many marketing researchers tend toward the latter argument, this study contends that conversation on Twitter provides valuable insight into the thoughts, actions and opinions of large and diverse populations. Indeed, even seemingly trivial tweets lend a unique perspective on the details of individuals' lives as they contain information on how individuals spend their days and how their moods change over time (Golder and Macy, 2011). Furthermore, the brevity of tweets and mobile-ready structure of Twitter itself offers the unique advantage of having a real-time perspective on how these factors change over time.

Presently, analysis of Twitter data focuses on the text of the tweets. This study utilizes other data encoded and/or displayed in the Twitter user's public profile, such as his or her pictures, geographic location, user ID, and the date and time each tweet was published. While some pieces of information provide helpful metadata when analyzing networks of Twitter users, others can offer key insights into the lives and characteristics of the users themselves. Of particular importance to those interested in gathering demographic data are the users' profile pictures – the primary photograph that the user chooses to represent himself/herself within Twitter. A preliminary search of 100 twitter profiles sampled for this study revealed that 75% of users have an identifiable face or full body shot as their profile picture. These pictures, which can be easily mined and stored by the researcher, provide the primary source of information for data collection methods outlined in this article. However, additional information that can be mined from the page, such as username or geotag location, additional uploaded photos and content of tweets, may also provide valuable insight for future research projects.

Who Uses Twitter?

As of December, 2012 there were 500 million registered Twitter users. According to the Pew Internet and American Life Project (2012), the percentage of Internet users who are on Twitter has doubled since November 2010 and as of 2012 stood at 16%. This population is dominated by younger individuals (i.e., those under the age of 50). African Americans internet users are more likely than whites or Hispanics internet users to use Twitter, as are urban dwelling internet users as opposed to internet users who live in rural or suburban areas. The same Pew study finds that 27% of internet users between ages 18 and 29 use Twitter, compared to 16% of internet users between the ages of 30 and 49, 10% of internet users between the ages of 50 and 64, and 2% of internet users over 65. Likewise, about 26% of African American internet users are on Twitter, compared with 14% of white, non-Hispanic Internet users and 19% of Hispanic internet users. Gender is approximately evenly distributed on Twitter; 17% of male internet users are on Twitter and 15% of female internet users also use Twitter.

Extracting Data from Twitter

Scraping data using the Twitter API

Web scraping has gained prominence among social science researchers as a means of collecting large amounts of data to explore topics such as election forecasting, tracking social trends, and time usage (Tumasjan 2010; Golder and Macy, 2011; Naaman et al. 2011). The term refers to the process of using an external computer program to extract data from a web platform – which is usually coded in HTML - and organize the data into a readable form. In order to automate communication with a web platform, the scraping program must obtain the platform's application programming interface (API), which is a standardized system of programming instructions that allows web platforms to access and share information from one another.[ii] In the same way that the web page's interface provides the user directives for interaction, the API helps guide communication between web programs. When applied to web scraping, the API allows the researcher to specify which elements of information he or she wishes to retrieve from the primary platform. Like many web tools, web platforms often release their API for researchers to use. API based commands are then embedded within an additional coding language – such as python or PHP as a means of refining the search to include specific keywords or queries.

Twitter maintains multiple options for accessing data. A common approach to collecting Twitter data involves collecting a small fraction of the entire volume of Twitter traffic. This approach, referred to as tapping the Twitter "firehose," produces massive amounts of data, with most of the data collected being unrelated to a given research question.

We describe an alternative approach based that targets the Twitter activity most related to a researcher's objective. Specifically, this study utilizes Twitter's rest API[iii] to collect all new tweets that match keyword search queries created within the past nine days. There are a number of user friendly interfaces available that allow researchers to run to API based code to scrape data from a given web page. Data for this project were collected using the free web scraping platform ScraperWiki.[iv] ScraperWiki is a collaborative online environment in which individuals develop and share scripts in Python, PHP and Ruby that are designed to collect and store online information from various websites.[v] Scraperwiki prevents the collection of duplicate data. In addition to this, the code used for this project was designed to prevent the collection of re-tweets, thus ensuring that the data collected is composed of original content from the Twitter users themselves.

One advantage of using Twitter's rest API is the ability to gather information using exact queries. This capacity allows the researcher to search for the specific attributes or behaviors of interest in a more precise manner than would be possible using the streaming API, which restricts searches to particular keywords rather than complete phrases. In this case, queries were designed to capture information on individuals who express a refusal to participate in the 2012 United States presidential election. Using the rest API, the researcher can also exclude individuals who disagree with a particular candidate (who might say "I'm not voting for Romney" for example) by excluding tweets containing words or phrases that reflect this phenomenon ("for Romney" or "Romney" in this example). We can also exclude many users who are discussing voting in other contexts (e.g., for a contestant on a television show) using keywords. Note that this exclusion element requires the researcher to familiarize him or herself with the nature of the behavior or characteristic at hand in order to develop a preliminary understanding of the terms necessary to exclude that are potentially related to voting that do not refer to the query at hand (in this case, a refusal to vote in the 2012 U.S. presidential election) such as “homecoming” or names of popular television shows. Exploring culturally, regionally, and timely appropriate means to construct queries is also important, and highlights the value of involving social scientists in the data collection process.

Due to the idiosyncratic and temporal nature of text information on Twitter, tweets were loosely monitored during the initial data collection process and some exclusion terms were added as they arose within the data. These irrelevant tweets – which compose the minority of the total body of data collected – were systematically removed later in the data processing step. The process used to clean this information will be discussed in the following paragraphs. A complete list of the queries and exclusion terms used is shown in *Table 1*.

[Insert Table 1 About Here]

Scrapers were run from October 16th, 2012 until November 9th, 2012. During this time, a total of 13,442 tweets were collected. From this pool we created a subfile of 500 randomly sampled working tweets from which we intended to gather demographic data. We used this sample to test multiple variations of survey tasks designed to collect demographic data about these users, the qualifications of the individuals intended to complete these tasks and the strategies used to remove irrelevant information from this data.

Cleaning Twitter Data

We use a two-pronged approach for retrieving relevant Twitter data. First we try to efficiently design queries and exclusion terms to retrieve the most accurate/relevant data (discussed above). Second, we filter or clean irrelevant tweets. We cleaned the data first by compiling lists of key, potentially irrelevant terms using familiarity with the content of the tweets as well as word frequency analysis using (1) Wordle[vi] and (2) detection using text mining tools. We then compared these cleaning techniques to hand coding performed by a research team member, which we assume to be the most accurate and complete filtering method used.

In our first cleaning strategy, we developed a list of key terms by first reading briefly through approximately 100 tweets in order to develop an understanding of how tweets were structured and which themes were prominent. In this case, terms related to race – such as “color” or “skin” or terms referring to women “she” or “her” were among those generally indicative of irrelevant information. Examples of irrelevant tweets include “I am not voting for maria cantwell because she voted yes for ndaa which is unacceptable” and “This is a Patriot!@kevlmar I am not voting my skin color, but voting for the future of my country #Military #Veteran #Heroes.”

We then used Wordle to create a list of keywords from a subsample of 500 tweets that might indicate irrelevant tweets to find and remove tweets including these terms. Wordle is a free online platform that creates word frequency clouds of text segments using a principle component analysis based algorithm and publishes these images to the web (www.wordle.com). Figure 1 provides our Wordle cloud. The size of the term in the Wordle, denotes the frequency at which this term was used, with larger words indicating higher frequency. As expected, the largest terms include “voting”, “vote”, and “refuse”. However, other large non-relevant terms can also be identified. For example, one particularly large term was “Casillas.” This is likely due to a large amount of tweeting regarding the fact that the Real Madrid goalkeeper, Iker Casillas was voting to decide the Ballon d`Or, the European Footballer of the Year award, which was occurring during the time of data collection. When we compared this technique to the hand coded data, we found that 47% of irrelevant tweets were cleaned by filtering based on the Wordle results.

[Figure 1 About Here]

In our second cleaning technique we replaced the use of Wordle with text mining techniques implemented using R's text mining package, *tm*[vii], as a means of finding terms that might signal irrelevant tweets. This package allows the user to organize the text content of a data file into a body of text called a *corpus,* and then display increasingly larger or more refined lists of the most frequently occurring words in the document. Reducing our word list to about 15 to 20 terms yielded a collection of terms similar to those visible within the Wordle. This technique also yielded the addition of "Christie" – in reference to Chris Christie, governor of New Jersey and apparent rumors of his potential candidacy in future presidential elections. When we compared this technique to the hand coded data, we found that 51% of irrelevant tweets were cleaned. This signals a marginal improvement over the use of Wordle – a tool that while easy to use is somewhat difficult to read – as a means of searching for potentially irrelevant terms within the data. Due to its somewhat preferable cleaning capacity, the data filtered using the R text analysis method are used for the results portion of this study.

Coding data from Twitter: Amazon Turkers

The following section describes Amazon's Mechanical Turk – a platform through which individuals can pay workers to perform short tasks for small fees – and the way in which this tool was used as a means of coding demographic data. This stage of the data collection process is perhaps the most important methodological contribution of this study. Crowdsourcing human intelligence is an essential step in extracting demographic information encoded as images rather than text. Below we discuss the details of the data collection procedure, as well as the ways in which Amazon's Mechanical Turk has been successfully used/implemented as a resource in previous studies.

Amazon's Mechanical Turk

Amazon's Mechanical Turk (AMT) platform is a marketplace for work that requires human rather than artificial intelligence. Within this platform, individuals, known as *requesters*, post brief tasks that can be performed in minutes in exchange for a dollar or less. These small assignments – called human intelligence tasks, or HITs – typically involve requests that are difficult or impossible for artificial intelligence to complete. Examples include tagging images, transcribing text from images, or answering questions about website content. Requesters have the ability to customize the price, format, and duration of their HITs, as well as set qualifications for the employees– or *Turkers* – who are permitted to view and/or complete these HITs. Turkers are anonymous, independent contractors who are identifiable only by their unique ID numbers. Each Turker's work history and overall approval rating is also available for view and can be used by requesters as a qualification for filtering. Despite their anonymity, some demographic information about the Turkers is known as the result of past survey research efforts. In addition to this, the survey instrument used for this study gathered administrative data about the Turkers themselves. The following paragraphs will address this.

Use of Amazon Turkers for Social Science Research

Previous studies have shown that Turkers can be highly reliable experimental research subjects (Mason and Suri 2010). It has been shown that Turkers behave and react similarly to research subjects within a laboratory setting and produce results of comparable quality (Bhurmester, Kwang and Gosling 2011; Mason and Suri 2010). Furthermore, using the AMT platform often allows experimental researchers to quickly and easily reach out to a larger, more stable and more diverse population than they might have otherwise been able to. In research experimenting directly with Turkers, Snow (2008) found that in regards to many language processing tasks such as affect determination, Turkers are just as effective as and less expensive than expert labelers. Marge and colleagues (2009) affirm the ability of Turkers to transcribe audio files; of the 20,116 words transcribed by the Turkers, only 997 (4.96%) contained errors. Urbano and colleagues (2010) asked Turkers to categorize pieces of music based on similarity, and again found that the Turkers performed as well as experts for a lesser price.

In addition to providing a successful platform for experimental research, the AMT also provides a valuable space for survey distribution. Researchers have expressed positive attitudes toward the potential accuracy and representativeness of the Turkers as survey subjects. Behrend (2011), for example, distributed a short survey to both the Turkers and a sample of university students as a means of comparing the psychometric properties of each. This study found that Turkers and university students behaved similarly and displayed similar judgment, but that the Turkers held a significant advantage for survey research in that they comprise a significantly more diverse respondent pool.

In terms of their demographic representation, Ipierotis (2010) finds that populations of Turkers are concentrated within two primary locations – approximately 50% are from the United States and 40% are from India. Turkers are overwhelmingly female (approximately 70%), and younger than the general population (51% of Turkers are between the ages of 21 and 35). Turkers also have a slightly lower yearly income than the general population of U.S. Internet users; over 60% of U.S. based Turkers have incomes below $60K. They also have small families (55% have no families). The Turkers sampled for this study generally parallel those surveyed by Ipierotis (2010). In addition, these Turkers are highly educated (44% have a bachelor's or master's degree). Many (44%) report that the AMT is their main source of income, although this trend is more representative of International Turkers rather than U.S. Turkers. Finally, the large number of HITs created for this study were completed by relatively few Turkers (N=48). The mean completion rate for the Turkers was 63 HITs with a range of 1 to 510 HITs. In addition, a large proportion of HITs (44%) were completed by Turkers with a bachelor's or master's degree.

Description of Methodology

In order to gather demographic information on the Twitter users who report a refusal to vote in the 2012 Presidential elections, Turkers were asked to view each user's profile picture and evaluate their sex, age, race, grooming and attractiveness. Categories for sex included male and female. For age, Turkers were asked to identify Twitter users according to both a numeric age range (from below 12 to 60+ years old), as well as a general age categories (child, adolescent, adult, senior). In order to test the accuracy and consistency of Tukers' evaluations of race – a particularly difficult survey question, as perceptions of race are shaped by culture - the survey procedure included three race questions that varied in terms of complexity and inclusiveness. The most basic included only categories for black and white with an option for "cannot tell." The somewhat more complex version of this question added a category for Asian. The most complex of the three included both Asian and Hispanic. Evaluations of attractiveness and grooming were both measured on a five point, ascending Likert- scale ranging from very unattractive/very poorly groomed to very attractive/very well groomed. These questions were drawn from the National Longitudinal Study of Adolescent Health[viii] which asked similar questions to the survey interviewers in all four waves of the study. Note that our questions differ slightly because we do not include a “don’t know” response option.

Also included in the survey were questions regarding the characteristics of the Turkers. Turkers were asked to state their sex, age, education level, the amount of time they spend per week on the AMT, and the whether the AMT provides their primary source of income. This metadata was compared to findings mentioned previously regarding the demographic composition of the Turkers by Ipierotis (2010), and as a means of confirming that the Turkers constitute a more demographically diverse respondent pool than traditional university samples and therefore might provide more reliable results when assessing socially constructed characteristics such as age category (i.e. child, adolescent, adult, senior) or race. The full survey instrument is included in *Appendix B.*

In regards to survey structure, this study administered a full questionnaire with all descriptive categories and Turker metadata questions mentioned in the previous paragraphs but used only the simplest (black/white) race evaluation question. This “full” survey contained a total of eleven questions. The two additional race questions were administered as separate surveys. Each completed full survey questionnaire yielded $0.10 for the Turker; each completed one question race evaluation yielded $0.02 cents for the Turker, resulting in an overall average hourly pay rate of $7.45 per hour.

In order to test the reliability of the Turkers’ evaluations, each photo was shown to three separate US and International Turkers. Consistency between these Turkers was monitored for each survey question. In addition to this, the same HITs were administered to both US based and international Turkers as a means of comparing reliability and results between the two groups (note that among both US and International Turkers three Turkers were asked to assess each photo). In order to ensure the most accurate results, only “master” Turkers – those who have completed at least 1000 approved HITs and have at least a 95% approval rating – were permitted to view or complete the HITs. Table 3 displays the reliability of the U.S. based, international and total Turker pool for each survey question.

These tables confirm previous suggestions from Behrend (2011) that Turkers provide a good respondent base for surveys. Overall, the Turkers prove to be very reliable in regards to their assessment of sex, age, age categorization and race. For all of these questions, the majority of responses were unanimous among the Turkers. The lowest total agreement rate for questions of this type is 52% (numeric age HITs assessed by International Turkers) and the highest is 83% (white/black race assessment HITs by U.S. based Turkers). Those that were not unanimous were generally agreed upon by two or more Turkers. For all questions, total disagreement is rare (between 1% and 5% of HITs for all questions pertaining to Age, Race and Sex). More categorically complex questions – such as those pertaining to numeric age or race including black, white, Asian and Hispanic – display slightly lower total agreement frequencies, although the rate at which two or more Turkers agree on these questions is still high (ranging from 95% to 100% for all Turkers). Subjective questions pertaining to attractiveness and grooming proved more difficult for the Turkers to consistently assess (total HIT agreement ranges from 21% to 29% for all Turkers). Nonetheless, for these questions there are relatively few cases in which no Turkers agreed; at most 24% of HITs for questions of this sort display no agreement. Finally, there seems to be little difference in the overall reliability of U.S. based and international Turkers.

3. RESULTS

Given our confidence in the reliability of the Turkers, we then consider the results of the Turkers’ evaluations of Twitter user profile pictures. The following tables display the demographic characteristics of the 500 Twitter users sampled for this study as determined by the Turkers. These tables are broken down to display evaluations for both US and International Turkers, as well as the results from the raw, quasi-auto filtered, and hand coded data. Note that for a Twitter user to be categorized in a particular way for any given question two or more Turkers had to agree upon that categorization. These results are shown in *Tables 4a-4h.*

Include *Tables 4a-4h* about here

According to these results, the majority of non-voters on Twitter are male (49.2% to 54.0%) adults between the ages of 19 and 35 (53.0% to 69.7%). When all racial categories are considered, the majority of non-voting Twitter users are reportedly white (44.4% to 48.7%), followed in descending frequency by black non-voters (26.4% to 29.7%), Hispanic non-voters (5.0% to 7.0%), and Asian non-voters (1.2% to 1.8%).

We also compare our results to data collected in a 2012 Pew Center report[ix]. This comparison is not intended to suggest that, given the current state of statistical modeling in this area, Twitter should be used to estimate population proportions or the sizes of certain populations. Instead, we currently find that the most compelling uses of Twitter data in studying real-time dynamics of social interactions, as we discuss below. We present this comparison as a means of evaluating our understanding of the differences between individuals on Twitter and those who are not. These results are shown in *Table 5*.

Include *Table 5* about here

As expected, the data presented in Tables 4 and 5 do not parallel the data on national non-voters gathered by Pew. It is clear that the Twitter estimates far exceed the Pew estimates in regards to the number of non-voters who are black and young. However, these inconsistencies are likely attributable to two factors. One, the population composition of Twitter does not align with the national population. A 2012 Pew Center Internet and American Life Project study of social media users[xi] revealed that the Twitter population is overrepresented by younger individuals (27% of internet users between the ages of 18 and 29 use Twitter as compared to 16% of users ages 30 to 49, 10% of users between ages 50 and 64 and 2% of users age 65 and older) and non-Hispanic black individuals (26% of non-Hispanic black individuals are on Twitter, as opposed to 14% of white non-Hispanic users and 19% of Hispanic users). The gender distribution of the Twitter population is relatively balanced (17% of male internet users and 15% of female internet users are Twitter users).

In addition to the demographic distribution of Twitter not aligning with the national population, the discrepancies between the demographic data on non-voters gathered in this study may be attributable to the effects of social desirability effects in previous surveys. According to Belli et al. (2013), many researchers believe that traditional surveys have a tendency to underrepresent the total number of non-voters, as non-voters often refuse to disclose this information for reasons of self-presentation. Given this systematic bias in survey design, it is possible that collecting information on deviant behaviors through Twitter – a space characterized by high levels of self-disclosure - provides *more* information on the individuals who engage in these behaviors than traditional surveys (Marwick and boyd, 2010). Future research will be required to determine if this might be the case.

When these discrepancies between the data collected from Twitter on those who refuse to vote and existing data on the non-voting population are considered, these estimates appear reasonable and support our understanding of the characteristics of individuals using Twitter.

3. DISCUSSION

It is becoming widely acknowledged that "social media offers us the opportunity for the first time to both observe human behavior and interaction in real time and on a global scale" (Golder and Macy, 2012: 7). Currently the majority of researchers who are taking advantage of social media data for social science research are not social scientists, but rather computer scientists and market researchers. Perhaps one reason for this trend is the fact that key pieces of information for sociological – and specifically demographic – research, such as age, race and gender, are difficult to extract from social media sites such as Twitter. Adding demographic information to Twitter data increases the breadth of social science research to which these data may be applied. Research could be done that examines not only collective attitudes and opinions, but also the composition of the groups driving these trends. This information could also be incorporated into network data and used as a means of examining the structure of groups that display deviant behaviors or opinions, as well as how this structure changes over time.

In this paper, we present a toolkit for extracting, processing, and analyzing data from Twitter. We believe that social media data, such as Twitter, present an opportunity for a fundamentally different approach to social science research. As with all new data collection, Twitter has certain limitations to overcome. Although the capacity of Twitter data parallels that of existing data collection, it does not replicate the results of these methods and poses new challenges. Our goal, however, is to address enough of these challenges to make Twitter an accessible resource for a larger fraction of social scientists and, in doing so, explore the contexts in which Twitter data has the greatest potential contribution in social science research.

Our preliminary analysis indicates that it is possible to collect demographic information on Twitter users using a combination of available technologies. Gathering raw data from Twitter using the website's API is simple, inexpensive and quick. Although only 500 tweets were used directly in this exploratory study, the scraping platform used in this study managed to collect a fairly large sample of individuals who report engaging in a deviant behavior (N=13,442). Obtaining evaluations of the Twitter user's demographic characteristics – including sex, age and race – using the AMT proved efficient and effective. Similar to previous research (Bhurmester, Kwang and Gosling, 2011; Behrend, 2011; Mason and Suri, 2010) our results indicate that Turkers are a reliable source for coding information and that we have access to a highly skilled and motivated collective of online workers. Although the data cleaning techniques suggested in this study requires further exploration, the results of these data collection efforts are nonetheless promising. The resultant demographic breakdown of non-voters presented in this paper do not parallel existing, national data on non-voters exactly, they nonetheless yield results that make sense given the demographic biases of the Twitter environment. We are confident that the data yielded through these methods could be used to develop more complex models for social analysis.

Advantages of Using Twitter for Demographic Data Collection

There are a number of advantages associated with the use of Twitter as source of data. To begin, Twitter data is abundant and easy to access. Among the approximately 500 million current registered Twitter users, approximately 88.2% are not protected, meaning that all published content is available for view to all web users. This published material is considered public data; Twitter users do not need to issue approval for researchers to use their profile information. Although laws regarding the use of Twitter information as public data may change in the future, social scientists have the opportunity to capitalize on the availability of Twitter data as pre-documented insight into the collective attitudes, opinions and behaviors of internet users. In addition to this advantage, micro-blogging websites such as Twitter are often updated multiple times per day, which allows the researcher to track opinions and actions as they emerge and develop. While traditional surveys accomplish a similar task, they are nonetheless time consuming and costly to administer and cannot provide the same minute-to-minute insight that Twitter data can.

Beyond availability, Twitter data is often easy and inexpensive to collect. There are a number of tools available that allow researchers to collect archived information within social media sites such as Twitter. The source used for this study is Scraperwiki, an open source platform for rest API based web scraping in which members can develop and share code to gather information from particular websites. Researchers may also use R – a free, collaborate software computing language and software environment used primarily for statistical and graphical analyses – to scrape web information using from online platforms using a streaming API.

Finally, Twitter provides ready access to certain populations that are difficult to access using other means, as Twitter users tend to disclose a great deal about their personal lives within this space. As discussed earlier, representation of self on Twitter is unique from other social networking platforms. Though users must sign in via a password protected web portal to post tweets, the majority of Twitter profiles (88.2%) are visible to all internet users. In addition to this, networks within Twitter are undirected and often contain a mixture of familiar and unfamiliar connections. Given these conditions, norms for disclosure on Twitter are ambiguous. Twitter users must maintain an online presence that is simultaneously polished and genuine (Marwick and boyd 2010). Occasionally these users utilize Twitter as a platform for unfiltered personal expression and admit to non-normative ideas and actions. In addition to individuals who express refusal to vote in the 2012 presidential election, preliminary analyses for this study also found a number of individuals who engage in deviant behaviors such as drunk driving or expressing racial slurs.

Although Twitter data is not suitable for all research questions, there are particularly interesting applications that may serve to expand our knowledge about social processes. There are also ways to leverage apparent weaknesses of Twitter for scientific purposes. For example, the open nature of Twitter, individuals can follow a profile without being a “friend” of the person tweeting, rather than being weakness, can, instead, allow us to examine the influence of weaker social network ties for behaviors and opinions.

While Twitter is not representative of the total US population, this does not negate the use of Twitter to examine social questions and for theory generation. In fact, the overrepresentation of African Americans and young adults on Twitter can be used to better understand populations that are often underrepresented in most surveys. In addition, similar to the case study approach, Twitter can be used for developing theoretical generalizations, if not statistically generalizable conclusions (Small 2009). In other words, although this data cannot be used to draw conclusion about the actions and behaviors of any population beyond that of Twitter users specifically, it can nonetheless be used to make statements about social processes in general.

In addition to this, Twitter allows researchers to examine the impact of short term events on behaviors and attitudes in ways we would not be able to do on such a large scale with a survey. These changes could presumably be analyzed on a scale as minute as week to week or day to day. Some researchers – such as De Loungueville, Smith and Luraschi (2009) - have addressed this capacity by using Twitter data as a means of tracking reactions to time sensitive events such as forest fire outbreaks. Nonetheless, adding demographic data to these analyses would expand possibilities for model building and the capacity for making predictions.

Challenges of Using Twitter for Demographic Data Collection

One major challenge associated with the use of Twitter data for social science research is the idiosyncratic nature of the data and need to remove irrelevant tweets from the data, as these may skew results. This paper attempts to provide a method for removing irrelevant tweets that allows the researcher to forego coding each tweet by hand. Hand coding can prove costly and time consuming for a research team, especially for projects that seek to utilize the "big data" type availability of Twitter information. Nonetheless, procedures proposed in this study catch at best approximately half of the total number of irrelevant tweets within the data and still require the researcher to develop some firsthand familiarity with the content of the tweets. Future research will address this challenge of data cleaning. Proposed methods may include subsetting and hand coding a portion of a larger body of data and using the demographic information garnered from this subset when building predictive models. It is important to note, however, that for this study results remained fairly consistent regardless of the filtering approach used (i.e., no irrelevant tweets removed; tweets removed using a semi-automated word search; irrelevant tweets removed after hand coding). See *Table 4* in the appendix for details evidence of this trend. In addition, the inclusions of social scientists in research teams with statisticians and computer scientists using Twitter and other social media data for social science research may be particularly helpful in the creation of theoretically-informed coding schemas and decisions rules that can inform new ways to filter complex social data.

In addition to handling the unpredictability of user-generated data, analyses that use Twitter data must be careful to consider issues of representation when interpreting results. It is important to state that these proposed data collections provide information about a very particular respondent pool: individuals who *report* not voting on Twitter. As indicated in the results portion of this paper, it is clear that Twitter users are not representative of the national population. Furthermore, these methods rely on voluntary information. The findings represent only those individuals who offer information about their voting intentions; it does not reflect individuals who did not vote *and* did not report these intentions or individuals who *claimed* to have no intention to vote by did nonetheless. Indeed, there is likely a group of individuals who are making false claims and may be providing erroneous profile information. Identifying these individuals/profiles will require the use of network information to better predict outcomes based on profiles where behaviors and characteristics can be more easily "verified". In addition, the targeted sampling strategy (collecting profiles based on specific behaviors/attitudes) we use reduces the likelihood we are drawing from fake accounts or avatars. To be sure, future research will be necessary to better handle these issues.

4. CONCLUSION

Twitter is arguably the largest observational study of human behavior to date. Not only is this source of data large and easily accessible by social scientists, we contend that there is tremendous opportunity for sociologists to use Twitter data for social science research but recognize that currently a barrier exists regarding the use of this data for demographic research. The purpose of this paper is to suggest a systematic and scalable means of gathering demographic data from Twitter – including age, race, and gender – as a means of overcoming this challenge. Supplementing textual data from Twitter with this additional information could open up brand new opportunities for social research and could allow demographers to model and predict behaviors and attitudes on a large scale and/or among difficult to reach populations. Restated, the potential for Twitter in social science research is yet to fully be articulated. However, we believe there are exciting opportunities to use this research to investigate social problems and other phenomena, but as a research community we cannot really explore these opportunities without (i) widespread access and familiarity with the data by social scientists and (ii) reliable information about demographic data. These are tremendous challenges, but overcoming them is worthwhile if doing so allows social science to play a role in utilizing one of the largest sources of social information available.

REFERENCES

Achrekar, Harshavardhan, Avinash Gandhe, Ross Lazarus, Ssu-Hsin Yu, and Benyuan Liu. 2011. "Predicting Flu Trends using Twitter data." in *First International Workshop on Cyber-Physical Networking Systems (CPNS) 2011, IEEE Infocom*. Shanghai, China.

Beevolve. 2012. An Exhaustive Study of Twitter Users Across the World. http://www.beevolve.com/twitter-statistics/. Retrieved March 1, 2013.

Belli, Robert F., Michael W. Traugott, Margaret Young, and Katherine A. McGonagle. 1999. ''Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring.'' *Public Opinion Quarterly* 63:90–108.

boyd, danah. 2009. "Twitter: “pointless babble” or peripheral awareness + social grooming? ." vol. 2012. http://www.zephoria.org/thoughts/archives/2009/08/16/twitter_pointle.html.

Brickman Bhutta, Christine. 2012. "Not by the Book." *Sociological Methods & Research* 41:57-88.

Conover, M., J. Ratkiewicz, M. Francisco, B. Gonçalves, A. Flammini, and F. Menczer. 2011. "Political polarization on twitter." in *5th International Conference on Weblogs and Social Media (ICWSM), 2011*, *Proceedings of the 5th International Conference on Weblogs and Social Media* Barcelona, Spain.

De Longueville, Bertrand De, RS Smith, and G Luraschi. 2009. “Omg, from here, i can see the flames!: a use case of mining location based social networks to acquire spatio-temporal data on forest fires.” … Based Social Networks (c):73–80. Retrieved March 9, 2013 (http://dl.acm.org/citation.cfm?id=1629907).

Diakopoulos, Nicholas A. and David A. Shamma. 2010. "Characterizing debate performance via aggregated twitter sentiment." Pp. 1195-1198 in *Proceedings of the 28th international conference on Human factors in computing systems*. Atlanta, Georgia, USA: ACM.

Goffman, Erving. 1959. *The Presentation of Self in Everyday Life* New York, NY: Doubleday.

Golder, Scott A. and Michael W. Macy. 2011. "Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures." *Science* 333:1878-1881.

Golder, Scott and Michael Macy. 2012. "Social Science with Social Media." *ASA Footnotes,* 40.

Heaivilin, N., B. Gerbert, J.E. Page, and J.L. Gibbs. 2011. "Public Health Surveillance of Dental Pain via Twitter." *Journal of Dental Research* 90:1047-1051.

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):377–386. Retrieved October 25, 2012 (http://bst.sagepub.com/cgi/doi/10.1177/0270467610385893).

Holbrook, Allyson L. and Jon A. Krosnick. 2010. "Social desirability bias in voter turnout reports: Tests using the item count technique." *Public Opinion Quarterly* 74:37-67.

Ipeirotis, Panagiotis G., Foster Provost, and Jing Wang. 2010. "Quality management on Amazon Mechanical Turk." Pp. 64-67 in *Proceedings of the ACM SIGKDD Workshop on Human Computation*. Washington DC: ACM.

Janus, Alexander L. 2010. "The Influence of Social Desirability Pressures on Expressed Immigration Attitudes*." *Social Science Quarterly* 91:928-946.

Java, Akshay, Xiaodan Song, Tim Finin, and Belle Tseng. 2007. "Why we twitter: understanding microblogging usage and communities." Pp. 56-65 in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. San Jose, California: ACM.

Krishnamurthy, Balachander, Phillipa Gill, and Martin Arlitt. 2008. "A few chirps about twitter." Pp. 19-24 in *Proceedings of the first workshop on Online social networks*. Seattle, WA, USA: ACM.

Lassen, David S. and Adam R. Brown. 2011. "Twitter: The Electoral Connection?" *Social Science Computer Review* 29:419-436.

Lowe, John B., Margaret Barnes, Cynthia Teo, and Stephanie Sutherns. 2012. "Investigating the use of social media to help women from going back to smoking post-partum." *Australian and New Zealand Journal of Public Health* 36:30-32.

Marge, M. , Banerjee, S., & Rudnicky, A. I. (2010). Using the Amazon Mechanical Turk for transcription of spoken language. In J. Hansen (Ed.), Proceedings of the 2010 IEEE Conference on Acoustics, Speech and Signal Processing (pp. 5270–5273). IEEE.

Marwick, Alice E. and danah boyd. 2010. "I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience." *New Media & Society*.

Moreno, Megan A., Allison Grant, Lauren Kacvinsky, Katie G. Egan, and Michael F. Fleming. 2012. "College Students' Alcohol Displays on Facebook: Intervention Considerations." *Journal of American College Health* 60:388-394.

Naaman, Mor, Hila Becker, and Luis Gravano. 2011. "Hip and trendy: Characterizing emerging trends on Twitter." *Journal of the American Society for Information Science and Technology* 62:902-918.

Pew Research Center. November 1, 2012. "Nonvoters: Who They Are, What They Think." Pew Research Center for the People & the Press. Retrieved February 23, 2013 (www.people-press.org/2012/11/01/nonvoters-who-they-are-what-they-think/ 4/).

Pew Research Center. February 14, 2013. "The Demographics of Social Media Users - 2012." Pew Internet and American Life Project. Retrieved February 14, 2013 (http://www.pewinternet.org/Reports/2013/Social-media-users.aspx)

Reips, Ulf-Dietrich and Pablo Garaizar. 2011. "Mining twitter: A source for psychological wisdom of the crowds." *Behavior Research Methods* 43:635-642.

Small, Mario Luis. 2009. Unanticipated Gains: Origins of Network Inequality in Everyday Life. Oxford: Oxford University Press.

Snow, R., O'Connor, B., Jurafsky, D., & Ng, A. Y. (2008). Cheap and fast—but is it good? Evaluating non-expert annotations for natural language tasks. In M. Lapata & H. T. Ng (Eds.), Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 254–263). New York, NY: ACM.

Tourangeau, Roger and Ting Yan. 2007. "Sensitive questions in surveys." *Psychological Bulletin* 133:859-883.

Tumasjan, A., T. O. Sprenger, P. G. Sandner, and I. M. Welpe. 2010. "Predicting elections with twitter: What 140 characters reveal about political sentiment." Pp. 178-185 in *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*.

Urbano, J., Morato, J., Marrero, M., & Martín, D. (2010). Crowd- sourcing preference judgments for evaluation of music similarity tasks. In M. Lease, V. Carvalho, & E. Yilmaz (Eds.), Proceedings of the ACM SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation (CSE 2010) (pp. 9–16). Geneva, Switzerland.

Valkenburg, Patti M., Jochen Peter, and Alexander P. Schouten. 2006. "Friend Networking Sites and Their Relationship to Adolescents' Well-Being and Social Self-Esteem " *CyberPsychology & Behavior* 9:584-590.

Yardi, Sarita and Danah Boyd. 2010. "Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter." *Bulletin of Science, Technology & Society* 30:316-327.

APPENDICIES

Appendix A: Figures and Tables

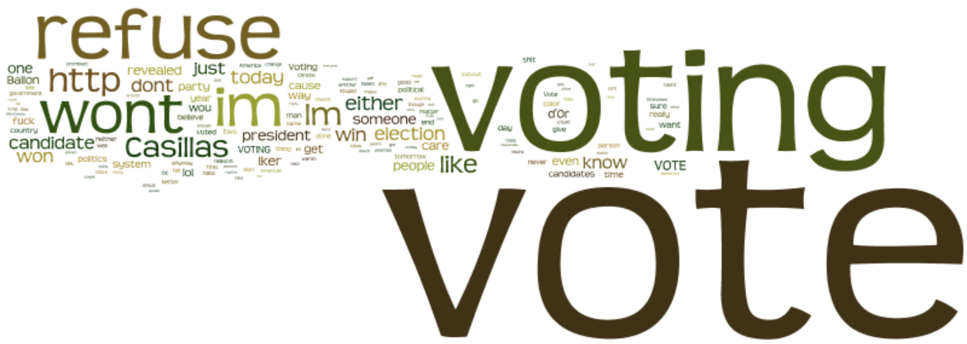

Figure 1: **Wordle Cloud.** This figure illustrates the use of Wordle for preliminary text analysis. Larger terms signify that the words occur more frequently within the document.

Table 1: **Not Voting Search Queries**

Query	# Results
"I am not voting"	1953
"I'm not voting"	4584
"I will not vote"	1966
"I won't vote"	784
"I am not going to vote"	200
"I'm not going to vote	239
"I'm not gonna vote"	160
"I am not gonna vote"	23
"I refuse to vote"	1150
"I don't plan to vote"	6
"I do not plan to vote"	2
"I didn't register to vote"	217
"I will never vote"	926
"I ain't voting"	995
"I ain't registered"	52
"I did not register to vote"	28
"I'll never vote"	157

Exclusion terms: -"EMA"-"AMA" –"Romney" -"Obama" –"xfactor" –"x-factor" –"x factor" –"#xfactor"'

Table 2a: **Turker Characteristics**

Question	**All Turkers (n=48)**	**US Turkers (n=26)**	**International Turkers (n=22)**
Main income source	44%	42%	45%
Education			
High School	15%	19%	9%
Some College	27%	35%	18%
Associate's Degree	13%	15%	9%
Bachelor's	27%	15%	41%
Master's Degree	17%	15%	18%
Age			0%
19 to 25	13%	0%	27%
26 to 35	56%	69%	41%
36 to 45	31%	31%	32%
Sex			0%
Male	40%	27%	55%
Female	60%	73%	45%

Table 2b: **Turker HIT Completion**

	All Turkers (n=48)	**US Turkers (n=26)**	**International Turkers (n=22)**
Mean Amount Completed	62.5	63.04	61.86
Hours on Turk Website			
1-2 Hours	4%	0%	9%
4-8 Hours	13%	8%	18%
8-20 Hours	31%	38%	23%
20-40 Hours	15%	19%	9%
40 Hours or more	38%	35%	41%

Table 3: **Turker reliability**

	% 3 agree	**% 2 of 3 agree**	**% none agree**
Age			
US (N=26)	56	40	4
International (N=22)	52	46	0
Total (N=48)	54	43	3
Age category			
US	64	31	4
International	60	38	2
Total	62	35	3
Race (white/black)			
US	83	14	3
International	79	18	3
Total	81	16	3
Race (including Asian)			
US	77	19	4
International	80	18	2
Total	79	18	3
Race (including Asian, Hispanic)			
US	74	22	4
International	74	21	5
Total	74	22	5
Sex			
US	84	14	2
International	83	16	1
Total	83	15	2
Attractiveness			
US	29	58	14
International	25	57	18
Total	27	58	16
Grooming			
US	26	61	14
International	21	55	24
Total	23	58	19

Notes:

Attractiveness and grooming based on a 5 point Likert Scale

***Table 4a-4h*: Demographic Composition of Non-Voters on Twitter (as evaluated by Amazon Turkers)**

Table 4a: Twitter User Sex

Turk Nation	Filter	Percent Male	Percent Female	Percent Cannot Tell	Percent No Agreement
All	Full	52.1 (48.4, 58.8)	43.7 (40.0, 47.3)	3.2 (1.9, 4.5)	1.0 (0.3, 1.7)
All	Partial	52.3 (49.0, 55.7)	42.9 (39.5, 46.2)	3.5 (2.3, 4.7)	1.3 (0.5, 2.0)
All	None	51.6 (48.5, 54.7)	42.0 (38.9, 45.1)	4.9 (3.6, 6.2)	1.5 (0.7, 2.3)
US	Full	53.1 (47.9, 58.3)	44.9 (47.9, 58.3)	0.6 (.00, 1.3)	1.4 (0.2, 2.6)
US	Partial	53.6 (48.9, 58.4)	44.0 (39.3, 48.7)	0.5 (0.0, 1.1)	1.9 (0.6, 3.2)
US	None	54.0 (49.6, 58.4)	43.0 (38.7, 47.3)	0.8 (0.0, 1.6)	2.2 (0.9, 3.5)
International	Full	51.1 (45.9, 56.3)	42.4 (37.3, 47.5)	5.9 (3.5, 8.3)	0.6 (0.0, 1.3)
International	Partial	51.1 (46.3, 55.8)	41.7 (37.0, 46.4)	6.6 (4.2, 8.9)	0.7 (0.0, 1.5)
International	None	49.2 (44.8, 53.6)	41.0 (36.7, 45.3)	9.0 (6.5, 11.5)	0.8 (0.0, 1.6)

Table 4b: Twitter User Numeric Age

Turk Nation	Filter	Percent Below 12	Percent 12 to 18	Percent 19 to 35	Percent 36 to 60
All	Full	1.0 (0.3, 1.7)	14.7 (12.1, 17.4)	64.0 (60.5, 67.6)	5.3 (3.7, 7.0)
All	Partial	1.3 (0.5, 2.0)	15.7 (13.3, 18.1)	62.8 (59.5, 66.0)	4.8 (3.4, 6.2)
All	None	1.5 (0.7, 2.3)	16.7 (14.4, 19.0)	58.3 (55.2, 61.4)	4.4 (3.1, 5.7)
US	Full	0.8 (0.0, 1.8)	9.3 (6.3, 12.3)	69.7 (64.9, 74.4)	4.8 (2.6, 7.0)
US	Partial	1.2 (0.2, 2.2)	10.3 (7.4, 13.2)	68.6 (64.2, 73.0)	4.2 (2.3, 6.1)
US	None	1.4 (0.4, 2.4)	11.6 (8.8, 14.4)	63.6 (59.4, 67.8)	3.8 (2.1, 5.5)
International	Full	1.1 (0.0, 2.2)	20.2 (16.1, 24.4)	58.4 (53.3, 63.5)	5.9 (3.5, 8.3)
International	Partial	1.4 (0.3, 2.5)	21.1 (17.2, 24.9)	56.9 (52.2, 61.6)	5.4 (3.2, 7.5)
International	None	1.6 (0.5, 2.7)	21.8 (18.2, 25.4)	53.0 (48.6, 57.4)	5.0 (3.1, 6.9)
Turk Nation	**Filter**	**Percent 60+**	**Percent Cannot Tell**	**Percent No Agreement**	
All	Full	0.1 (0.0, 0.4)	11.7 (9.3, 14.0)	3.1 (1.8, 4.4)	
All	Partial	0.1 (0.0, 0.3)	12.6 (10.4, 14.9)	2.7 (1.6, 3.8)	
All	None	0.1 (0.0, 0.3)	16.2 (13.9, 18.5)	2.8 (1.8, 3.8)	
US	Full	0.0 (No CI)	11.5 (8.2, 14.8)	3.9 (1.9, 6.0)	
US	Partial	0.0 (No CI)	12.4 (9.3, 15.5)	3.3 (1.6, 5.0)	
US	None	0.0 (No CI)	16.0 (12.8, 19.2)	3.6 (2.0, 5.2)	
International	Full	0.3 (0.0, 0.8)	11.8 (8.5, 15.1)	2.2 (0.7, 3.8)	
International	Partial	0.2 (0.0, 0.7)	12.9 (9.7, 16.1)	2.1 (0.7, 3.5)	
International	None	0.2 (0.0, 0.6)	16.4 (13.2, 19.6)	2.0 (0.8, 3.2)	

Table 4c: Twitter User Age Category

Turk Nation	Filter	Percent Child	Percent Adolescent	Percent Adult
All	Full	1.0 (0.3, 1.7)	14.6 (12.0, 17.2)	70.8 (67.4, 74.1)
All	Partial	1.3 (0.5, 2.0)	15.6 (13.1, 18.0)	68.9 (65.7, 72.0)
All	None	1.5 (0.7, 2.3)	16.6 (14.3, 18.9)	63.9 (60.9, 66.9)
US	Full	0.8 (0.0, 1.8)	9.3 (6.3, 12.3)	75.8 (71.4, 80.3)
US	Partial	1.2 (0.2, 2.2)	10.3 (7.4, 13.2)	74.0 (69.8, 78.2)
US	None	1.4 (0.4, 2.4)	11.6 (8.8, 14.4)	68.8 (64.7, 72.9)
International	Full	1.1 (0.0, 2.2)	19.9 (15.8, 24.1)	65.7 (60.8, 70.7)
International	Partial	1.4 (0.3, 2.5)	20.8 (17.0, 24.7)	63.7 (59.1, 68.3)
International	None	1.6 (0.5, 2.7)	21.6 (18.0, 25.2)	59.0 (54.7, 63.3)
Turk Nation	**Filter**	**Percent Senior**	**Percent Cannot Tell**	**Percent No Agreement**
All	Full	0.1 (0.0, 0.4)	10.4 (08.2, 12.6)	3.1 (1.8, 4.4)
All	Partial	0.1 (0.0, 0.3)	11.5 (09.3, 13.6)	2.7 (1.6, 3.8)
All	None	0.1 (0.0, 0.3)	14.9 (12.7, 17.1)	3.0 (1.9, 4.1)
US	Full	0.0 (No CI)	9.3 (6.3, 12.3)	4.8 (2.6, 7.0)
US	Partial	0.0 (No CI)	10.5 (07.6, 13.5)	4.0 (2.1, 5.8)
US	None	0.0 (No CI)	13.8 (10.8, 16.8)	4.4 (2.6, 6.2)
International	Full	0.3 (0.0, 0.8)	11.5 (08.2, 14.8)	1.4 (0.2, 2.6)
International	Partial	0.2 (0.0, 0.7)	12.4 (09.3, 15.5)	1.4 (0.3, 2.5)
International	None	0.2 (0.0, 0.6)	16.0 (12.8, 19.2)	1.6 (0.5, 2.7)

Table 4d Twitter User Attractiveness

Turk Nation	Filter	Percent Unattractive	Percent About Average	Percent Attractive
All	Full	3.7 (2.3, 5.0)	27.5 (24.2, 30.8)	33.7 (30.2, 37.2)
All	Partial	3.3 (2.1, 4.5)	26.8 (23.8, 29.8)	33.7 (30.6, 36.9)
All	None	3.0 (1.9, 4.1)	26.1 (23.6, 28.8)	31.9 (29.0, 34.8)
US	Full	1.7 (0.3, 3.0)	26.1 (21.6, 30.7)	39.6 (34.5, 44.7)
US	Partial	1.6 (0.4, 2.8)	24.8 (20.7, 28.9)	39.8 (35.2, 44.5)
US	None	1.4 (0.4, 2.4)	23.8 (20.1, 27.5)	38.6 (34.3, 42.9)
International	Full	5.6 (3.2, 8.0)	28.9 (24.2, 33.6)	27.8 (23.2, 32.5)
International	Partial	4.9 (2.9, 7.0)	28.8 (24.5, 33.1)	27.6 (23.4, 31.9)
International	None	4.6 (2.8, 6.4)	28.4 (24.4, 32.4)	25.2 (21.4, 29.0)
Turk Nation	**Filter**	**Percent Very Attractive**	**Percent Cannot Tell**	**Percent No Agreement**
All	Full	3.9 (2.5, 5.4)	14.5 (11.9, 17.1)	16.7 (14.0, 19.5)
All	Partial	4.0 (2.7, 5.3)	15.6 (13.1, 18.0)	16.6 (14.1, 14.9)
All	None	3.8 (2.6, 5.0)	19.3 (16.9, 21.7)	15.9 (13.6, 18.2)
US	Full	3.7 (1.7, 5.6)	14.9 (11.2, 18.6)	14.0 (10.4, 17.7)
US	Partial	3.7 (1.9, 5.5)	15.7 (12.2, 19.1)	14.3 (11.0, 17.6)
US	None	3.2 (1.7, 4.7)	19.4 (15.9, 22.9)	13.6 (10.6, 16.6)
International	Full	4.2 (2.1, 6.3)	14.0 (10.4, 17.7)	19.4 (15.3, 23.5)
International	Partial	4.2 (2.3, 6.1)	15.5 (12.0, 18.9)	19.0 (15.3, 22.7)
International	None	4.4 (2.6, 6.2)	19.2 (15.7, 22.7)	18.2 (14.8, 21.6)

Table 4e: Twitter User Grooming

Turk Nation	Filter	Percent Poorly Groomed	Percent About Average	Percent Well Groomed
All	Full	4.6 (3.1, 6.2)	25.6 (22.4, 28.8)	35.5 (32.0, 39.0)
All	Partial	4.2 (2.9, 5.6)	24.7 (21.8, 27.6)	34.4 (31.2, 37.6)
All	None	4.0 (2.8, 5.2)	24.2 (21.5, 26.9)	32.6 (29.7, 35.5)
US	Full	2.2 (0.7, 16.1)	27.0 (22.4, 31.6)	44.4 (39.2, 49.5)
US	Partial	2.3 (0.9, 3.8)	24.8 (20.7, 28.9)	43.8 (39.1, 48.5)
US	None	2.2 (0.9, 3.5)	24.0 (20.3, 27.7)	41.8 (37.5, 46.1)
International	Full	7.0 (4.4, 9.7)	24.2 (19.7, 28.6)	26.7 (22.1, 31.3)
International	Partial	6.1 (3.8, 8.4)	24.6 (20.5, 28.7)	25.1 (20.9, 29.2)
International	None	5.8 (3.8, 7.8)	24.4 (20.6, 28.2)	23.4 (19.7, 27.1)
Turk Nation	**Filter**	**Percent Very Well Groomed**	**Percent Cannot Tell**	**Percent No Agreement**
All	Full	1.4 (0.5, 2.3)	13.5 (11.0, 16.0)	19.4 (16.5, 22.3)
All	Partial	2.2 (1.2, 3.2)	14.6 (12.3, 17.0)	19.8 (17.1, 22.5)
All	None	2.1 (1.2, 3.0)	18.1 (15.7, 20.5)	19.0 (16.6, 21.4)
US	Full	0.8 (0.0, 1.8)	12.6 (9.2, 16.1)	12.9 (9.4, 16.4)
US	Partial	1.4 (0.3, 2.5)	13.6 (10.3, 16.8)	14.1 (10.8, 17.3)
US	None	1.2 (0.2, 2.2)	17.0 (13.7, 20.3)	13.8 (10.8, 16.8)
International	Full	2.0 (0.5, 3.4)	14.3 (10.7, 18.0)	25.8 (21.3, 30.4)
International	Partial	3.0 (1.4, 4.7)	15.7 (12.2, 19.1)	25.5 (21.4, 29.7)
International	None	3.0 (1.5, 4.5)	19.2 (15.7, 22.7)	24.2 (20.4, 28.0)

Table 4f: Twitter User Race (Black/White only)

Turk Nation	Filter	Percent White	Percent Black	Percent Cannot Tell	Percent No Agreement
All	Full	56.2 (52.5, 59.8)	29.8 (26.4, 33.1)	11.8 (9.4, 14.2)	2.2 (1.2, 3.3)
All	Partial	54.7 (51.3, 58.0)	30.0 (26.9, 33.0)	12.9 (10.6, 15.1)	2.5 (1.4, 3.5)
All	None	52.7 (49.6, 55.8)	28.6 (25.8, 31.4)	16.2 (13.9, 18.5)	2.5 (1.5, 3.5)
US	Full	56.5 (51.3, 61.6)	29.8 (25.0, 34.5)	11.2 (8.0, 14.5)	2.5 (0.9, 4.2)
US	Partial	54.8 (50.1, 59.5)	30.2 (25.9, 34.6)	12.4 (9.3, 15.5)	2.6 (1.1, 4.1)
US	None	53.2 (48.8, 57.6)	29.0 (25.0, 33.0)	15.4 (12.2, 18.6)	2.4 (1.1, 3.7)
International	Full	55.9 (50.7, 61.1)	29.8 (25.0, 34.5)	12.4 (8.9, 15.8)	2.0 (0.5, 3.4)
International	Partial	54.6 (49.8, 59.3)	29.7 (25.4, 34.1)	13.3 (10.1, 16.6)	2.3 (0.9, 3.8)
International	None	52.2 (47.8, 56.6)	28.2 (24.3, 32.1)	17.0 (13.7, 20.3)	2.6 (1.2, 4.0)

Table 4g: Twitter User Race(with Asian)

Turk Nation	Filter	Percent White	Percent Black	Percent Asian
All	Full	49.9 (46.2, 53.5)	29.2 (25.9, 32.6)	2.1 (1.1, 3.2)
All	Partial	49.1 (45.7, 52.4)	29.5 (26.4, 32.6)	2.3 (1.3, 3.4)
All	None	46.9 (43.8, 50.0)	27.8 (25.0, 30.6)	2.6 (1.6, 3.6)
US	Full	50.3 (45.1, 55.5)	29.2 (24.5, 33.9)	2.5 (0.9, 4.2)
US	Partial	49.4 (44.7, 54.2)	29.5 (25.2, 33.8)	2.8 (1.2, 4.4)
US	None	47.2 (42.8, 51.6)	28.0 (24.1, 31.9)	3.2 (1.9, 4.7)
International	Full	49.4 (44.2, 54.6)	29.2 (24.5, 33.9)	1.7 (0.3, 3.0)
International	Partial	48.7 (44.0, 53.5)	29.5 (25.2, 33.8)	1.9 (0.6, 3.2)
International	None	46.6 (42.2, 51.0)	27.6 (23.7, 31.5)	2.0 (0.8, 3.2)
Turk Nation	**Filter**	**Percent Other**	**Percent Cannot Tell**	**Percent No Agreement**
All	Full	3.7 (2.3, 5.0)	11.5 (9.2, 13.9)	3.7 (2.3, 5.0)
All	Partial	3.4 (2.2, 4.6)	12.5 (10.3, 14.7)	3.2 (2.0, 4.3)
All	None	3.3 (2.2, 4.4)	16.3 (14.0, 18.6)	3.1 (2.0, 4.2)
US	Full	0.6 (0.0, 1.3)	12.4 (8.9, 15.8)	5.1 (2.8, 7.3)
US	Partial	0.7 (0.0, 1.5)	13.1 (9.9, 16.3)	4.4 (2.5, 6.4)
US	None	0.6 (0.0, 1.3)	16.8 (13.5, 20.1)	4.2 (2.4, 6.0)
International	Full	6.7 (4.1, 9.3)	10.7 (7.5, 13.9)	2.2 (0.7, 3.8)
International	Partial	6.1 (3.8, 15.0)	11.9 (8.9, 15.0)	1.9 (0.6, 3.2)
International	None	6.0 (3.9, 8.1)	15.8 (12.6, 19.0)	2.0 (0.8, 3.2)

Table 4h: Twitter User Race (with Asian, Hispanic)

Turk Nation	Filter	Percent White	Percent Black	Percent Asian	Percent Hispanic
All	Full	47.5 (42.3, 52.7)	28.7 (24.0, 33.3)	1.4 (0.2, 2.6)	5.6 (3.2, 8.0)
All	Partial	47.7 (44.3, 51.0)	28.9 (25.9, 32.0)	1.4 (0.6, 2.2)	6.1 (4.5, 7.7)
All	None	45.5 (42.4, 48.6)	27.3 (24.5, 30.1)	1.6 (0.8, 2.4)	5.7 (4.3, 7.1)
US	Full	46.6 (42.2, 51.0)	28.2 (24.3, 32.1)	1.4 (0.4, 2.4)	6.4 (4.3, 8.5)
US	Partial	48.7 (44.0, 53.5)	29.7 (25.4, 34.1)	1.2 (0.2, 2.2)	7.0 (4.6, 9.4)
US	None	46.6 (42.2, 51.0)	28.2 (24.3, 32.1)	1.4 (0.4, 2.4)	6.4 (4.3, 8.5)
International	Full	47.5 (42.3, 54.7)	28.7 (24.0, 33.3)	1.4 (0.2, 2.6)	5.6 (3.2, 8.0)
International	Partial	46.6 (41.9, 51.3)	28.1 (23.8, 32.4)	1.6 (0.4, 2.8)	5.2 (3.1, 7.2)
International	None	44.4 (40.0, 48.8)	26.4 (22.5, 30.3)	1.8 (0.6, 3.0)	5.0 (3.1, 6.9)
Turk Nation	**Filter**	**Percent Other**	**Percent Cannot Tell**	**Percent No Agreement**	
All	Full	0.0 (No CI)	12.4 (8.9, 15.8)	4.5 (2.3, 6.6)	
All	Partial	0.5 (0.0, 0.9)	10.9 (8.8, 13.0)	4.6 (3.2, 6.0)	
All	None	0.8 (0.2, 1.4)	4.5 (4.3, 7.1)	4.6 (3.3, 5.9)	
US	Full	1.2 (0.2, 2.2)	12.2 (9.3, 15.1)	4.0 (2.3, 5.7)	
US	Partial	0.9 (0.0, 1.9)	8.7 (6.0, 11.3)	3.7 (1.9, 5.5)	
US	None	1.2 (0.2, 2.2)	12.2 (9.3, 15.1)	4.0 (2.3, 5.7)	
International	Full	0.0 (No CI)	12.4 (8.9, 15.8)	4.5 (2.3, 6.6)	
International	Partial	0.0 (No CI)	13.1 (9.9, 16.3)	5.4 (3.2, 7.5)	
International	None	0.4 (0.0, 1.0)	16.8 (13.5, 20.1)	5.2 (3.3, 7.1)	

Table 5: **Pew Institute Data on Non-Voters**

Sex	
Men	52%
Women	48%
Race/Ethnicity	
White, non-Hispanic	59%
Black, non-Hispanic	10%
Hispanic	21%
Age	
18-29	36%
30-49	35%
50-64	20%
65+	8%

Appendix B: Survey Instrument

Look at the Twitter profile picture and identify the following characteristics of the main person in the picture:

1. What is the sex of the main person in the picture?
   a. Make
   b. Female
   c. Cannot tell
2. Given the following choices, what is the race of the main person in the picture?
   a. White
   b. Black
   c. Cannot tell
3. Given the following choices, what is the race of the main person in the picture? *(Question distributed as separate survey)*
   a. White
   b. Black
   c. Asian
   d. Cannot tell
4. Given the following choices, what is the race of the main person in the picture? *(Question distributed as separate survey)*
   a. White
   b. Black
   c. Asian
   d. Hispanic
   e. Cannot tell
5. Given the following choices, what is the approximate age of the main person in the picture?
   a. Below 12 years
   b. 13 to 18 years
   c. 19 to 35 years
   d. 36 to 60 years
   e. 60+ years
   f. Cannot tell
6. What is the age category of the main person in the picture?
   a. Child
   b. Adolescent/teenage
   c. Adult
   d. Senior
   e. Cannot tell
7. How attractive is the main person in the picture?
   a. Very unattractive
   b. Unattractive
   c. Attractive
   d. Very attractive
   e. Cannot tell
8. How well groomed is the main person in the picture?
   a. Very poorly groomed
   b. Poorly groomed
   c. About Average

d. Well groomed
e. Very well groomed
f. Cannot tell

Finally, tell us a little bit about yourself:

1. What is your sex?
    a. Male
    b. Female
2. What is your age
    a. Under 18
    b. 19 to 25
    c. 26 to 35
    d. 36 to 45
    e. 46 to 55
    f. 56 to 65
    g. Over 65
3. What is your highest level of education?
    a. Some high school
    b. High school
    c. Some college
    d. Associates degree
    e. Bachelors degree
    f. Graduate degree, Masters
    g. Graduate degree, Doctorate
4. How many hours *per week* do you spend on the Mechanical Turk?
    a. Less than one hour
    b. 1 to 2 ours
    c. 2 to 4 hours
    d. 4 to 8 hours
    e. 8 to 20 hours
    f. 20 to 40 hours
    g. 40 hours or more
5. Is the Mechanical Turk your main source of income?
    a. Yes
    b. No

ENDNOTES

[i] See https://www.mturk.com/mturk/welcome for more detailed information.

[ii] See http://dev.twitter.com for more detailed information.

[iii] Twitter also maintains a streaming API, which provides a continuous stream of Tweets and user information based on single word queries.

[iv] See https://scraperwiki.com/ for more detailed information.

---

[v] The scraper used in this project was loosely drawn from a scraper created by Scraperwiki user by Paul Bradshaw designed to collect data on the 2011 Balham riots.

[vi] See http://www.wordle.net/ for more detailed information.

[vii] See http://cran.r-project.org/web/packages/tm/index.html for more detailed information.

[viii] See http://www.cpc.unc.edu/projects/addhealth for more detailed information.

[ix] The analysis in this report is based on telephone interviews conducted among national samples of adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia. Three surveys are referenced in this report: October 24-28, 2012 (2,008 adults); October 4-7, 2012 (1,511 adults); and September 12-16, 2012 (3,019 adults).

[xi] Data from this study come from a national survey conducted between November 14 and December 9, 2012 on landline and cell phones and in English and in Spanish. The results reported here come from the 1,802 respondents who are Internet users and the margin of error is +/- 2.6 percentage points.