

Characterizing Geographic Variation in Well-Being using Tweets

H. Andrew Schwartz,* Johannes C. Eichstaedt,* Margaret L. Kern, Lukasz Dziurzynski, Megha Agrawal
Gregory J. Park, Shrinidhi K. Lakshmikanth, Sneha Jha, Martin E. P. Seligman, and Lyle Ungar

University of Pennsylvania

hansens@seas.upenn.edu, jeich@sas.upenn.edu

Richard E. Lucas

Michigan State University

Abstract

The language used in tweets from 1,300 different US counties was found to be predictive of the subjective well-being of people living in those counties as measured by representative surveys. Topics, sets of co-occurring words derived from the tweets using LDA, improved accuracy in predicting life satisfaction over and above standard demographic and socio-economic controls (age, gender, ethnicity, income, and education). The LDA topics provide a greater behavioural and conceptual resolution into life satisfaction than the broad socio-economic and demographic variables. For example, tied in with the psychological literature, words relating to outdoor activities, spiritual meaning, exercise, and good jobs correlate with increased life satisfaction, while words signifying disengagement like 'bored' and 'tired' show a negative association.

Introduction

Social media has proven to be remarkably useful for tracking geographical variations in health. Google uses search queries to measure trends in flu, providing earlier indication of disease spread than the "gold standard" data from the Centers of Disease Control (CDC), which is based on hospital reports (Ginsberg et al. 2008). Similarly, tweets show the variation in allergies by region or time of year (Paul and Dredze 2011).

This paper addresses another related health issue, subjective well-being, as measured by life satisfaction (LS). As we explain below, the importance of life satisfaction goes far beyond the obvious attraction of positive emotion; it contributes to health, productivity, and other positive life outcomes (Pressman and Cohen 2005). From the standpoint of social media research, studying life satisfaction offers novel issues beyond those in predicting flu or allergies, as our goals include not just *predicting* regional variation in happiness, but also in *understanding* the factors contributing to it. The use of the language in tweets gives us insights into some of those factors.

We collected a billion tweets from June 2009 to March 2010, mapped as many as possible to the US counties that

they were sent from, and correlated the words used in the tweets (in the form of LDA-generated *word topics*) with life satisfaction, as measured by questionnaires answered in those counties. We also have demographic information (age, sex, ethnicity) and indicators of socio-economic status (income and education) by county, which we used as controls in a predictive model. We find that word use gives additional predictive accuracy above the socio-demographic controls in predicting LS. Lastly, and toward *understanding*, we show visualizations of the *word topics* that predict LS; these provide informative intuitions about what factors the model is capturing.

In the following we provide some basic background on subjective well-being before turning to a more detailed description of our analysis method and results.

Subjective Well-being and its measurement

Happiness matters. For example, when a sample of Britons were asked what the prime objective of their government should be – "greatest happiness" or "greatest wealth", 81% answered with happiness (Easton 2006). In a set of other studies conducted around the world, 69% of people on average rate well-being as their more important life outcome (Diener 2000). Psychologists still argue about how happiness should be defined, but few would deny that people desire it.

Governments around the world are starting to put more effort into measuring subjective well-being in their countries, moving beyond the common economic-based indicators such as Gross Domestic Product (Stiglitz, Sen, and Fitoussi 2009b; 2009a). Surveys by organizations such as Gallup, and government agencies (e.g., the CDC in the US) increasingly are including one or more subjective well-being questions in their questionnaires. However, survey research is expensive, in terms of time and resources. We would like to find faster and cheaper methods to assess well-being. Further, it is important to not only measure well-being, but also attend to factors that contribute to, support, and improve it (Abramovitz, Scitovsky, and Inkeles 1973; Layard 2005; Seligman 2011).

Subjective well-being is not just an end in itself; there is increasing evidence that well-being improves multiple life domains, including health and immunity (Howell, Kern, and Lyubomirsky 2007), cardiovascular disease (Boehm and

Kubzansky 2012), career outcomes and social relationships (Lyubomirsky, King, and Diener 2005; Pressman and Cohen 2005), as well as longevity (Diener and Chan 2011). For example, college students who reported higher levels of cheerfulness also had higher incomes, greater job satisfaction, and less unemployment 19 years later (Diener et al. 2002).

Subjective well-being refers to how people evaluate their lives in terms of cognition (i.e., satisfaction with life) and emotion (positive and negative emotion). Most of the happiness research in social media has focused on the emotion component (i.e., sentiment analysis). For example, classifying the emotional affinity of sentences and characterizing happiness as a specific emotion are often the goals of semantic analysis of text (Alm, Roth, and Sproat 2005; Mihalcea and Liu 2006). Quercia et al. (2012) found Twitter users expressing positive (or negative) emotion to cluster together. Another study viewed happiness as a lack of deprivation, which is an index on socio-economic factors like income, education, health, crime, and employment (Quercia, Séaghdha, and Crowcroft 2012)—factors that may relate to subjective well-being but do not constitute a direct measure. O'Connor et al. (2010) were able to predict opinion polls based on sentiment analysis of tweets containing topical keywords. Others have looked at variation in positive and negative emotion word use in tweets, Facebook, or blogs across time (Dodds et al. 2011; Kramer 2010) or latitude (Dodds and Danforth 2010). Such methods, though useful for other goals, do not capture many of the nuances of subjective well being.

Here, we specifically focus on the cognitive-based evaluation of overall life satisfaction (LS), a broader evaluation of well-being than emotion alone provides. Our goal here is not to count positive sentiment words, but study the language of well-being, to better understand the multiple components that contribute to it.

Some of the above studies, particularly those attempting to track “happiness”, are not based on a “ground truth”, as measured, for example, in questionnaires; they look at variation in the use of an *ex ante* list of indicative words. In contrast, we find the words that correlate with life satisfaction as measured in questionnaires; this allows us to empirically construct far richer lexica associated with the many aspects of happiness.

LS has been widely studied in the psychological literature (Diener et al. 1999) and has been tracked by the CDC through their Behavioral Risk Factor Surveillance System, as well as by numerous countries around the world; the OECD has recently established authoritative guidelines for the measurement of subjective well-being (OECD 2013). It is assessed by asking people to respond to questions such as “In general, how satisfied are you with your life?”, with responses ranging from “very dissatisfied” to “very satisfied” (Diener et al. 1985). The response to such simple questions has provided useful comparisons of well-being both within and between nations (Diener et al. 2010). Although LS is a single indicator, it is influenced by many important areas of life such as having sufficient food and shelter, good relationships, and having the freedom to choose one’s daily activities (Diener et al. 1985). Which factors in particular

emerge as dominant contributors to LS is a matter of debate. Drawing on our massive social media dataset, we allow the data to tell their story about the most predictive human concerns and behaviors.

Geolocation and Twitter

We are not the first to make use of geolocation information in Twitter. Others have studied how word use in Twitter varies with location, and used word frequencies in twitter to predict geolocation (Eisenstein et al. 2010; Han, Cook, and Baldwin 2012; Hong et al. 2012; Cheng, Caverlee, and Lee 2010). Furthermore, Hecht et al. (2011) found people that people do not always provide real information in their location field. We keep in this in mind when mapping location fields to counties.

More closely related to this paper, as mentioned above, twitter messages have been analyzed to identify a range of health-related terms such as symptoms, syndromes and treatments to highlight geographical patterns in syndrome surveillance. Paul and Dredze developed a variant of topic models that captures the symptoms and possible treatments for ailments, traumatic injuries and allergies, discussed on Twitter, with a focus on general public health. They also explore the geographical patterns in the prevalence of such ailments (Paul and Dredze 2011).

Method

The primary goal of our method is to find language that characterizes subjective well-being over counties as measured by random life satisfaction phone surveys (Lawless and Lucas 2011). Here, we describe how we find the language that characterizes locational well-being, our approach to prediction, and how we map tweets to counties.

Differential Language Analysis

In working toward finding language characterizing well-being, we focus on lexical / topical features which are easily interpreted:

- **lexica:** hand-built lists of words including those from the psychological tool Linguistic Inquiry and Word Count (*LIWC*) (Pennebaker et al. 2007), as well as a terms associated with the *PERMA* (positive emotion, engagement, relationships, meaning in life, and accomplishment) construct of well-being (Seligman 2011). Each list of words is associated with a semantic or syntactic category, such as positive emotion, leisure, engagement, or pronouns. Usage is measured as the percentage of a county’s words which were within the given category.
- **topics:** clusters of lexico-semantically related words as derived automatically from Latent Dirichlet Allocation (LDA). We used social-media specific topics, 2000 in all, described in Schwartz et al. (2013), which were derived from 18 million Facebook status updates. To measure topic usage, we use the distributions provided by LDA such that a given county’s *topic* usage is defined as:

$$p(\text{topic}|\text{county}) = \sum_{\text{word} \in \text{topic}} p(\text{topic}|\text{word}) * p(\text{word}|\text{county})$$

where $p(\text{word}|\text{county})$ is the normalized word use by that *county* and $p(\text{topic}|\text{word})$, the probability of the topic given the word, is provided by LDA. Furthermore, we use the joint probability, $p(\text{word}, \text{topic})$, in order to determine a word's prevalence in a topic (i.e., when producing the tag cloud visualizations). While the hand-built *lexica* provide a theory-driven set of language, LDA gives us open-ended clusters of words from actual distributions in social media. It is a means to find unexpected characteristics of "happy counties".

We use a tweet-specific tokenizer (Gimpel et al. 2010) to identify words. Since the dataset is quite large, we use MapReduce of the tokenizer over a small hadoop cluster in order to aggregate words by counties. We then extract the category and topic features as described above.

After extraction of features, we run a correlation analysis between all categories and topics and the LS scores per county. We use ordinary least squares linear regression over standardized variables, which produces a Pearson r as correlation size. All results deemed positively or negatively correlated must pass a Bonferroni-corrected p-value of 0.05 (because we look at 2000 different topics, $p < (.05/2000)$).

Lastly, we found effective visualization of the results of such an analysis to be of critical importance since we essentially end up with 2000 correlations. We use word clouds, as demonstrated in Figure 2 to represent topics, where the size of the word within a cloud indicates the relative prevalence of the word ($p(\text{word}, \text{topic})$) within the topic. We only consider those topics deemed significant and plot the most highly correlated.

Predictive Models

For our predictive model of subjective well-being, we use the same features as differential language analysis, *lexica* and *topics*. After acquiring usage information for both types of features, we apply a log transform to reduce the variance (this reduces the effect language use outliers can have on model fit).

Controls: To compare with some of the best predictors of county-level well-being, we drew on demographic and socio-economic status (SES) from the U.S. Census Bureau. This included the following demographics from the 2010 US census:

- *median age*
- *sex* (percentage female)
- *minorities* (percentage black and Hispanic).

as well as estimates of the following SES variables from 2009:

- *median household income* (log-transformed)
- *educational attainment* (percentage high school graduates or higher; percentage bachelor's degrees or higher).

The two variables of educational attainment were combined into an overall educational attainment index by averaging the respective standardized scores across the counties. We refer to these as *controls*, seeking to discover whether our language models can add information beyond what these variables already contribute.

The *lexica*, *topics*, and *controls* are run through a LASSO (L^1 penalized) linear regression with life satisfaction over counties. Regularization via the L^1 penalty, which drives less-predictive features to be weighted zero (Tibshirani 1996), is preferred since our sample size is often smaller than the number of features.

Data Set and Geolocation

Our collection of tweets were from a random 10% collected between November 2008 and January 2010 via the the Twitter "garden hose". We use the 1,293 counties out of the 2,528 counties studied in (Lawless and Lucas 2011) for which we had at least 30,000 tweeted words (i.e., at least 1,000 words from each of 30 people). Additionally, we averaged LS over the years 2009 and 2010 to reduce measurement error (LS tends to be fairly stable over time (inter-year correlations around $r=.85$); variations due to measurement error are an issue in the smaller counties with reduced sample sizes, and averaging across years reduces this error).

Mapping tweets to counties is non-trivial (Hecht et al. 2011). Only a small fraction of the tweets have geolocation coordinates that can be mapped directly to counties. Instead, we rely on parsing the free-response location field that accompanies a tweet, sometimes containing a city/state pair or an individual city by itself, or sometimes non-identifiable phrases (e.g., "My Hous").

We use a cascaded set of rules to map location strings to US counties, designed with the goal of avoiding false positives (incorrect mappings) at the expense of finding fewer total mappings. After tokenizing the strings, we attempt to match country names with the tokens, starting from the right, and only keeping the messages that either mention the country as the United States (in one of several forms) or do not mention the country at all. Next, using the tokens preceding the country (if available), we attempt to match city and state names. When only cities can be matched, we use a table of city populations¹ and map a city to a state if it has a 90% likelihood of being in the particular state according to the population size of all cities of that name. We throw out cities for which we cannot determine city and state information following these criteria, as well as city names that also happen to be one of the 100 largest non-US cities.

In a random selection of 100 city and state pairs extracted from location strings of tweets which also had latitude and longitude coordinates, 84% mapped to the city associated with the coordinates. However, these coordinates often reflect where the user tweeted from rather than where they live. We therefore selected another random 100 city and state pairs, and judged 93% to have correctly identified the intended location. This evaluation is limited to assuming that those indicating valid city/state pairs are being honest.² These cities are then mapped onto the counties in which they lie, since our demographic and LS data are available at the county level.

¹downloaded here: <http://www.uscitieslist.org/>

²Note that should our mapping have many errors, this would make our task of predicting well-being by location more difficult.

data	<i>r</i>
Lexica	0.264
Topics	0.307
Topics & Lexica	0.307
Controls	0.435
Controls, Topics & Lexica	0.535

Table 1: Test set prediction accuracy of LS using LIWC and well-being word categories ('Lexica'), LDA Topics ('Topics'), demographic and SES variables ('Controls'), and all three together, as measured using the Pearson *r*.

After reducing to counties for which we have controls, LS, and which wrote at least 30,000 words, we are left with approximately 82 million county-mapped tweets. Correlating the number of tweets per county against the population of those counties ($r = .74$) suggests modest demographic representativeness despite a known twitter sampling bias in favor of urban and young populations (Hargittai and Litt 2011; Smith and Brenner 2012).

Results and Discussion

Prediction. The accuracy with which we can predict LS using different feature sets is shown in Table 1. The 1,293 counties for which we had at least 30,000 tweets word were randomly divided into 75% training / development examples (970 counties) and the remaining 25% (323 counties) held-out for testing. As is common in the social sciences, we use Pearson's *r* between the predicted life satisfaction scores and those measured from the survey to evaluate each model's predictive value. The reader should note that Pearson correlations between behavior (i.e., language use) and psychologically based variables rarely exceeds an *r* of 0.4 (Meyer et al. 2001).

The demographic and SES *controls* (age, sex, ethnicity, income, education) are more predictive than LDA-generated twitter word *topics* alone, which are, in turn more useful than the *lexica*, hand-crafted lists of words from LIWC and well-being theory. (See Appendix.) All three feature sets combined give significantly more accurate results than the controls alone, confirming that the words in tweets contain information beyond those in the control variables. We also see that the *lexica* do not seem to add additional predictive value when used in addition to the topics. However, we kept them in our analyses as they are useful in characterizing well-being, and have been used by others for predicting psychological variables (Kramer 2010; Sumner et al. 2012).

We've found both that language alone (*topic* and *lexica*) is predictive and that it contributes information above and beyond standard *controls*. One might find these results more surprising when considering that the people writing the tweets are unlikely to be the same as those sampled for the survey data. Other studies have suggested that happiness is contagious (Fowler and Christakis 2008), and our results seem to imply that one's well-being can be reflected by the words of a community sample.

Differential Language Analysis. Figure 1 shows regional variation in LS (a) as determined by survey data and (b) as predicted by our approach using 10-fold cross-validation. As shown above, the *topics* and *lexica* significantly improve prediction accuracy. More importantly, they also provide potential insight into what is being captured by the demographic and SES variables. Figure 2 shows the top 10 topics most highly correlated with LS.

The automatically-derived LDA topics include a remarkable range of cognitive and behavioral elements that have been empirically associated with subjective well-being. These include the following (labels in italics refer to the topics in Figure 2):

- *training, class, session, gym*: Across age and gender, physical activity relates to psychological well-being (Biddle and Ekkekakis 2006; Blumenthal and Gullette 2002). People report feeling better after exercising, and some evidence suggests that it reduces risk of depression and can act as an alternative to drug treatments in treating depression and other disorders (Lawlor and Hopker 2001; Mutrie 2004; Stathopoulou et al. 2006)
- *ideas, suggestions*: highlights the language of people tapping their social network for ideas, suggestions, opinions, and advice. Actively seeking counsel through social ties fits into a broader pattern of problem-centered coping and broadened thought-action repertoires exhibited by those with high levels of well-being (Folkman and Moskowitz 2000; Fredrickson and Joiner 2002). One tweet states, "So nice to to have ideas to think about after a week of meetings and writing reports about stuff we've already done."
- *money, support, donate*: A wide variety of pro-social activities have been shown to increase life satisfaction, including giving money and engaging in political activism (Klar and Kasser 2009; Dunn, Aknin, and Norton 2008). One tweet states, "...Buckwheat is on the list next time i decide to donate 2 Whole Foods."
- *meeting, conference*: Several of the topics tap various aspects of engagement. Engagement is a multidimensional construct that can be considered an important predictor or an indicator of well-being (Appleton, Christenson, and Furlong 2008; Fredricks et al. 2011). Engagement is considered a key part of healthy aging (Rowe and Kahn 1987) and is included as a core component of well-being in some theories (Seligman 2011). The construct includes forms that are psychological (e.g., fully concentrated and happily engrossed in activities, Schaufeli, Bakker, and Salanova, 2006; flow, Csikszentmihalyi, 1997), cognitive (e.g., valuing activities, self-regulation, goal-setting; Appleton et al., 2006); and behavioral (e.g., involvement, dedication; organizational citizenship behavior, Appleton et al., 2006) forms. The topic shown suggests communal engagement, perhaps in school or work groups - pointing to a behavioral type of engagement.
- *human, beings, nature, spiritual*: These words suggest that the connection to something larger than oneself is an important determinant of psychological well-being,

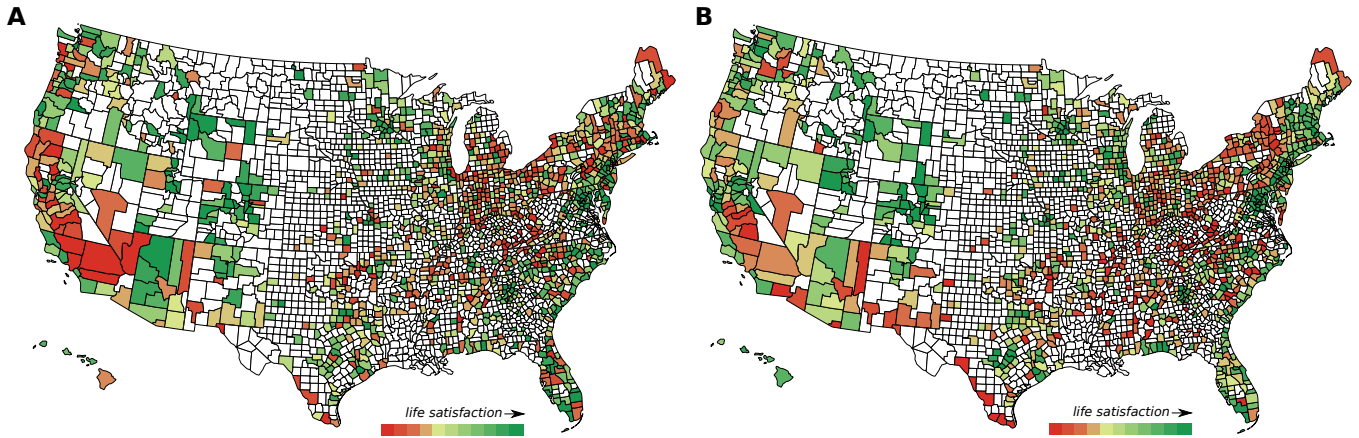


Figure 1: Map of the US showing life satisfaction (LS) as measured (A) using survey data and (B) as predicted using our combined model (controls + word *topics* and *lexica*). Green regions have higher satisfaction, while red have lower. White regions are those for which the language sample or survey size is too small to have valid measurements. (No counties in Alaska met criteria for inclusion; $r = 0.535$, $p < 0.001$)

aligned with recent psychological theory (Seligman 2011; Forgeard et al. 2011). In particular, the topics suggests a connection of all of humankind, and the concomitant “spiritual emotion” of compassion (Vaillant 2008; Armstrong 2010). One of the many proposed pathways of being connected in this way to more robust well-being is the increased ability to cope with life’s serious stressors, to maintain mental health in the face of adversity (Pargament 2001).

- *skills, management, business, learning*: LS is high in high value creation occupations in which continuous learning, roles of responsibility and skill development are valued. In congruence with this finding, previous research regarding county level life satisfaction has revealed moderate positive correlations between LS and employment in the “professional” occupation sector (as opposed to “construction”, “sales” or “service”) (Lawless and Lucas 2011). Furthermore, also consistent with this topic, it has been proposed that having a high percentage of employment in a county in the “creative” and “super-creative class” is positively associated with well-being (Rentfrow, Mellander, and Florida 2009).
- *experience, bound, wonderful*: This cluster represents engagement with life. LS is increased by experiencing life fully, adapting to both the good and bad (Frederick and Loewenstein 1999).

Space precludes showing the rest of the LS-positive topics, but of particular interest are a set of three related to the outdoors: sea/water, mountains, and hiking (see Figure 3). Counties with ocean or mountains tend to have more educated and wealthy populations, but also are strongly associated with recreation topics, which has been suggested to have an effect on happiness (Hartig, Mang, and Evans

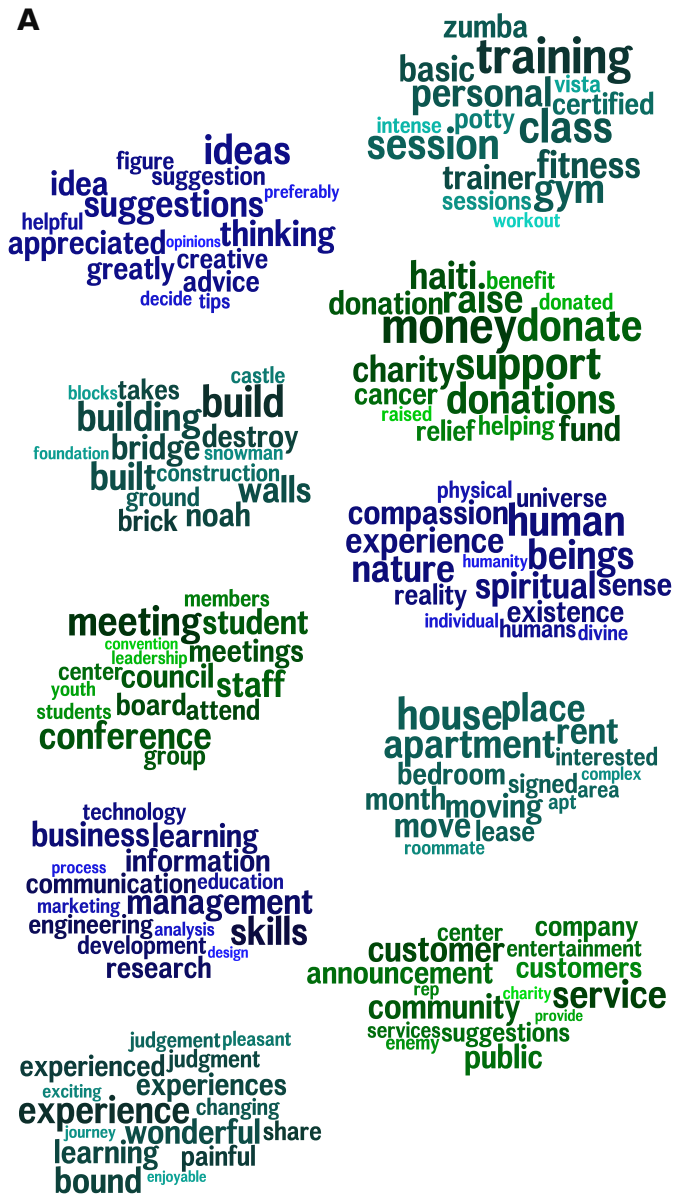
1991). Other positive topics (not shown³) include learning, ideas, money, meetings, ability, house-related terms, groups, computers and opportunities.

Negative topics (Figure 2B) are far less varied. They include fewer substantive terms, and more words relating to attitude. Some are explicitly negative (‘sick’, ‘hate’), but many are more indicative of disengagement: ‘bored’, ‘chill’, ‘wtf’. For example, one tweet states “... feel ur pain. sitting here with kids is all i ever do. bored out of my mind!”. These negative words are, of course, associated more strongly with younger people. This is suggestive of the empirical observation that older people tend, on average, to have higher subjective well-being, referred to by psychologists as the “aging positivity effect” (Carstensen and Mikels 2005).

A similar story is borne out by the dictionaries that psychologists have assembled to code text for different salient psychological dimensions (Linguistic Inquiry and Word Count, LIWC) (Pennebaker, Francis, and Booth 2001; Pennebaker et al. 2007). Table 2 shows that lexicon categories most predictive of positive LS are “money”, “work” and words tied to “achievement”, which supports the psychological theory that the experience of “accomplishment” is one of the pillars of subjective well-being (Seligman 2011). The use of plural personal pronouns such as “we” and “our” which we take to be proxies for a communal, pro-social orientation are highly correlated with the presence of LS, whereas “I” and “my” are highly correlated with its absence. This again supports conceptions of subjective well-being being supported by a focus on relationships (Seligman 2011), a construct sometimes alternatively referred to as “relatedness” (Ryan and Deci 2000). In terms of a priori lexica, the single strongest predictor of LS is a dictionary of words psychologists have assembled which they take to express disengagement (including words such as “sleepy”,

³see wwbp.org for full list of topics.

A



B



Figure 2: Top ten topics most positively correlated with well-being (A), and top two topics negatively correlated with well-being (B). Word size corresponds to prevalence within the topics. Topics are significantly correlated with LS at a Bonferroni-corrected $p < 0.001$.

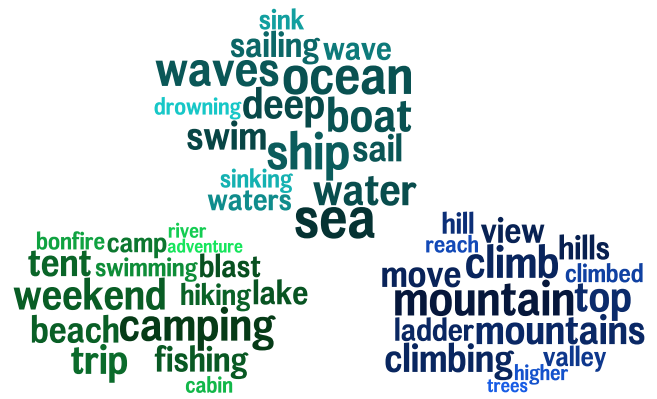


Figure 3: Three outdoor / nature activity topics positively correlated with well-being. Word size corresponds to prevalence within the topics. Topics are significantly correlated with LS at a Bonferroni-corrected $p < 0.001$.

“tired”, “bored”), and conversely, words associated with the “engagement” (‘excited’) construct emerge as some of the strongest positive predictors. In regard to (dis-)engagement, curated dictionaries and LDA topics tell the same story. Sets of words tied to positive and negative emotions (Pennebaker et al. 2007) show the predicted correlations with life satisfaction; the same is true for the LIWC subdictionaries which represent swear words and expressions of anger.

Conclusions

In looking at the correlation of word use with subjective well being, we found many patterns that have been observed in the well-being literature, including positive effects of pro-social activities, exercise, engagement at school and work, and openness to and engagement with life.

Words of disengagement in a county predicts lower life satisfaction. In addition to the LIWC and well-being hand-curated disengagement-related categories, we found many word topics of disengagement, often with different slang usages. Such words are also used more by younger people, who are on average less happy, but we suspect are more saliently used by more disengaged youth.

The word topics we found do often correlate with demographic and socio-economic status (SES) variables, but the words provide potential insight into what may underlie the correlation of these variables with LS. We know that happiness is roughly linear in the log of income (Deaton and Hestorr 2010), but it is interesting that it is donating money and having rewarding jobs that people in happier communities talk about, far more than what one can buy with the money. We argue that our methodology can offer insight on what specific aspects of people’s everyday experiences impact their life satisfaction. These specific pathways assist policy makers and psychologists to design effective interventions and evaluate the specific impact of policy decisions.

The fundamental result of this paper is perhaps surprising: we can predict (on average) the happiness of one set of

category	r
money (LIWC)	0.151 ***
work (LIWC)	0.145 ***
engagement (WB)	0.136 ***
positive emotion (WB)	0.127 ***
we (LIWC)	0.126 ***
positive emotion (LIWC)	0.124 ***
achievement (LIWC)	0.117 **
space (LIWC)	0.111 **
accomplishment (WB)	0.106 **
article (LIWC)	0.104 *
body (LIWC)	-0.134 ***
negative relationships (WB)	-0.135 ***
sexual (LIWC)	-0.137 ***
assent (LIWC)	-0.142 ***
shehe (LIWC)	-0.142 ***
swear (LIWC)	-0.146 ***
negative emotion (LIWC)	-0.152 ***
first-pers pron (LIWC)	-0.157 ***
anger (LIWC)	-0.160 ***
disengagement (WB)	-0.188 ***

Table 2: Top positively and negatively correlated categories from LIWC and our theory-driven well-being (WB) *lexica* along with Pearson correlation (r). (Bonferonni-corrected $p < 0.05^*$, 0.01^{**} , 0.001^{***})

people (those who answered the LS questionnaires) from the tweets of *other* people (people in the same county). This is, however, consistent with findings from other methodologies. People in the same county tend to share the same culture and environmental affordances (e.g., hiking, music, or good employment), and attitudes towards them (being excited or bored).

Happiness is asserted to be contagious (Fowler and Christakis 2008) and it has been suggested that although educated people are happier, on average, than less educated ones, there is an even stronger benefit to living in a community of educated people with arts, culture and entertainment (Lawless and Lucas 2011). Thus, the tweets of other people can indicate what it's like to live around them, influencing one's own happiness.

Disentangling the various causes and correlations is difficult. The CDC actively follows Google Flu trends to supplement its own tracking, but Google's algorithms are still being tweaked as they occasionally over or under-estimate flu rates due to popular media sources influencing search patterns (Butler 2013). Though such temporal issues do not directly apply when observing tweets over locations, analogous prediction noise likely occurs, requiring consideration within models like ours. Our work is a step towards a social media-based well-being predictor which uses a richer feature set than the simple hedonic measures in previous work. We hope this translates into the ability to estimate and better understand the subjective well-being of large populations with nuanced conceptual and behavioral resolution.

Acknowledgements

Support for this research was provided by the Robert Wood Johnson Foundation's Pioneer Portfolio, through a grant to Martin Seligman, "Exploring Concepts of Positive Health".

Appendix

We use handcrafted sets of words from two collections, PERMA (Seligman 2011), and LIWC (Pennebaker et al. 2007).

The PERMA lexicon is a collection of words relating to five components of positive psychology:

- *Positive emotion* (aglow, awesome, bliss . . .),
- *Engagement* (absorbed, attentive, busy . . .),
- *Relationships* (admiring, agreeable . . .),
- *Meaning* (aspire, belong . . .) and
- *Achievement* (accomplish, achieve, attain . . .).

For each of these five categories, we have both positive words – ones that connote, for example, *achievement*, and negative words, for example, *un-achievement* (amateurish, blundering, bungling . . .); or *engagement* and *disengagement* (bored, distracted, numb, sleepy . . .).

LIWC is a much broader collection of word categories, including everything from positive and negative emotion to swear words to parts of speech. There is a vast amount of social science research utilizing it (Pennebaker et al. 2007).

References

- Abramovitz, M.; Scitovsky, T.; and Inkeles, A. 1973. Economic growth and its discontents. *Bulletin of the American Academy of Arts and Sciences* 11–27.
- Alm, C.; Roth, D.; and Sproat, R. 2005. Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of the conference on Empirical Methods in Natural Language Processing*, 579–586. Association for Computational Linguistics.
- Appleton, J. J.; Christenson, S. L.; Kim, D.; and Reschly, A. L. 2006. Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology* 44(5):427–445.
- Appleton, J. J.; Christenson, S. L.; and Furlong, M. J. 2008. Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools* 45(5):369–386.
- Armstrong, K. 2010. *Twelve steps to a compassionate life*. Knopf Canada.
- Biddle, S. J., and Ekkekakis, P. 2006. Physically active lifestyles and well-being. In Huppert, F. A.; Baylis, N.; and Keverne, B., eds., *The science of well-being*. Oxford University Press. 141–170.
- Blumenthal, J. A., and Gullele, E. C. D. 2002. Exercise interventions and aging: Psychological and physical health benefits in older adults. In Schaie, K. W.; Levanthal, H.; and Willis, S. L., eds., *Effective health behavior in older adults: Societal impact on aging*. New York: Springer. 157–177.

- Boehm, J. K., and Kubzansky, L. D. 2012. The heart's content: The association between positive psychological well-being and cardiovascular health. *Psychological Bulletin* 138:655–691.
- Butler, D. 2013. When google got flu wrong. *Nature* 494(7436):155–156.
- Carstensen, L. L., and Mikels, J. A. 2005. At the intersection of emotion and cognition aging and the positivity effect. *Current Directions in Psychological Science* 14(3):117–121.
- Cheng, Z.; Caverlee, J.; and Lee, K. 2010. You are where you tweet: a content-based approach to geo-locating twitter users. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, 759–768. ACM.
- Csikszentmihalyi, M. 1997. *Creativity: Flow and the psychology of discovery and invention*. Harper Perennial.
- Deaton, A., and Hestorr, A. 2010. Understanding ppps and ppp-based national accounts. *American economic journal. Macroeconomics* 2(4):1–35.
- Diener, E., and Chan, M. Y. 2011. Happy people live longer: Subjective well-being contributes to health and longevity. *Applied Psychology: Health and Well-Being* 3(1):1–43.
- Diener, E.; Emmons, R. A.; Larsen, R. J.; and Griffin, S. 1985. The satisfaction with life scale. *Journal of personality assessment* 49(1):71–75.
- Diener, E.; Suh, E. M.; Lucas, R. E.; and Smith, H. L. 1999. Subjective well-being: Three decades of progress. *Psychological bulletin* 125(2):276.
- Diener, E.; Nickerson, C.; Lucas, R. E.; and Sandvik, E. 2002. Dispositional affect and job outcomes. *Social Indicators Research* 59(3):229–259.
- Diener, E.; Ng, W.; Harter, J.; and Arora, R. 2010. Wealth and happiness across the world: material prosperity predicts life evaluation, whereas psychosocial prosperity predicts positive feeling. *Journal of Personality and Social Psychology* 99(1):52.
- Diener, E. 2000. Subjective well-being: The science of happiness and a proposal for a national index. *American psychologist* 55(1):34.
- Dodds, P., and Danforth, C. 2010. Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. *Journal of Happiness Studies* 11(4):441–456.
- Dodds, P. S.; Harris, K. D.; Kloumann, I. M.; Bliss, C. A.; and Danforth, C. M. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PLoS ONE* 6(12):26.
- Dunn, E. W.; Aknin, L. B.; and Norton, M. I. 2008. Spending money on others promotes happiness. *Science* 319(5870):1687–1688.
- Easton, M. 2006. Britain's happiness in decline. *BBC News* 2.
- Eisenstein, J.; O'Connor, B.; Smith, N.; and Xing, E. 2010. A latent variable model for geographic lexical variation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 1277–1287. Association for Computational Linguistics.
- Folkman, S., and Moskowitz, J. T. 2000. Positive affect and the other side of coping. *American psychologist* 55(6):647–654.
- Forgeard, M. J. C.; Jayawickreme, E.; Kern, M.; and Seligman, M. E. P. 2011. Measuring wellbeing for public policy. *International Journal of Wellbeing* 1(1):49–106.
- Fowler, J. H., and Christakis, N. A. 2008. The dynamic spread of happiness in a large social network. *BMJ: British medical journal* 337:a2338.
- Frederick, S., and Loewenstein, G. 1999. Hedonic adaptation.
- Fredricks, J.; McColskey, W.; Meli, J.; Montrosse, B.; Mordica, J.; and Mooney, K. 2011. Measuring student engagement in upper elementary through high school: A description of 21 instruments. *Issues & Answers Report, REL* 98(098).
- Fredrickson, B. L., and Joiner, T. 2002. Positive emotions trigger upward spirals toward emotional well-being. *Psychological science* 13(2):172–175.
- Gimpel, K.; Schneider, N.; O'Connor, B.; Das, D.; Mills, D.; Eisenstein, J.; Heilman, M.; Yogatama, D.; Flanigan, J.; and Smith, N. A. 2010. Part-of-speech tagging for twitter: Annotation, features, and experiments. Technical report, DTIC Document.
- Ginsberg, J.; Mohebbi, M. H.; Patel, R. S.; Brammer, L.; Smolinski, M. S.; and Brilliant, L. 2008. Detecting influenza epidemics using search engine query data. *Nature* 457(7232):1012–1014.
- Han, B.; Cook, P.; and Baldwin, T. 2012. Geolocation prediction in social media data by finding location indicative words. In *Proceedings of COLING 2012*, 1045–1062. Mumbai, India: The COLING 2012 Organizing Committee.
- Hargittai, E., and Litt, E. 2011. The tweet smell of celebrity success: Explaining variation in twitter adoption among a diverse group of young adults. *new media & society* 13(5):824–842.
- Hartig, T.; Mang, M.; and Evans, G. W. 1991. Restorative effects of natural environment experiences. *Environment and behavior* 23(1):3–26.
- Hecht, B.; Hong, L.; Suh, B.; and Chi, E. H. 2011. Tweets from justin bieber's heart: the dynamics of the location field in user profiles. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, 237–246. ACM.
- Hong, L.; Ahmed, A.; Gurumurthy, S.; Smola, A. J.; and Tsioutsouliklis, K. 2012. Discovering geographical topics in the twitter stream. In *Proceedings of the 21st international conference on World Wide Web*, 769–778. ACM.
- Howell, R. T.; Kern, M. L.; and Lyubomirsky, S. 2007. Health benefits: Meta-analytically determining the impact of well-being on objective health outcomes. *Health Psychology Review* 1(1):83–136.
- Klar, M., and Kasser, T. 2009. Some benefits of being an

- activist: Measuring activism and its role in psychological well-being. *Political Psychology* 30(5):755–777.
- Kramer, A. 2010. An unobtrusive behavioral model of gross national happiness. In *Proc of the 28th int conf on Human factors in comp sys*, 287–290. ACM.
- Lawless, N. M., and Lucas, R. E. 2011. Predictors of regional well-being: a county level analysis. *Social Indicators Research* 101(3):341–357.
- Lawlor, D. A., and Hopker, S. W. 2001. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *Bmj* 322(7289):763.
- Layard, R. 2005. Rethinking public economics: The implications of rivalry and habit. *Economics and Happiness* 147–170.
- Lyubomirsky, S.; King, L.; and Diener, E. 2005. The benefits of frequent positive affect: does happiness lead to success? *Psychological Bulletin* 131(6):803.
- Meyer, G.; Finn, S.; Eyde, L.; Kay, G.; Moreland, K.; Dies, R.; Eisman, E.; Kubiszyn, T.; and Read, G. 2001. Psychological testing and psychological assessment: A review of evidence and issues. *American psychologist* 56(2):128.
- Mihalcea, R., and Liu, H. 2006. A corpus-based approach to finding happiness. In *Proceedings of the AAAI Spring Symposium on Computational Approaches to Weblogs*, 19.
- Mutrie, N. & Faulkner, G. 2004. Physical activity: Positive psychology in motion. In Linley, P. A., and Joseph, S., eds., *Positive psychology in practice*. New Jersey: John Wiley and Sons. 146–164.
- OECD. 2013. Oecd guidelines on measuring subjective well-being”.
- OConnor, B.; Balasubramanian, R.; Routledge, B. R.; and Smith, N. A. 2010. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*, 122–129.
- Pargament, K. I. 2001. *The psychology of religion and coping: Theory, research, practice*. Guilford Press.
- Paul, M. J., and Dredze, M. 2011. You are what you tweet: Analyzing twitter for public health. In *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*.
- Pennebaker, J. W.; Chung, C. K.; Ireland, M.; Gonzales, A.; and Booth, R. J. 2007. The development and psychometric properties of liwc2007 the university of texas at austin. *LIWC.NET* 1:1–22.
- Pennebaker, J. W.; Francis, M. E.; and Booth, R. J. 2001. Linguistic inquiry and word count: Liwc 2001. *Word J. Of The Int Ling Assoc*.
- Pressman, S. D., and Cohen, S. 2005. Does positive affect influence health? *Psychological Bulletin* 131(6):925–971.
- Quercia, D.; Capra, L.; and Crowcroft, J. 2012. The social world of twitter: Topics, geography, and emotions. In *The 6th international AAAI Conference on weblogs and social media, Dublin*.
- Quercia, D.; Séaghdha, D.; and Crowcroft, J. 2012. Talk of the city: Our tweets, our community happiness. *Proc. of AAAI ICWSM*.
- Rentfrow, P. J.; Mellander, C.; and Florida, R. 2009. Happy states of america: A state-level analysis of psychological, economic, and social well-being. *Journal of Research in Personality* 43(6):1073–1082.
- Rowe, J. W., and Kahn, R. L. 1987. Human aging: usual and successful. *Science*.
- Ryan, R. M., and Deci, E. L. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55(1):68.
- Schaufeli, W. B.; Bakker, A. B.; and Salanova, M. 2006. The measurement of work engagement with a short questionnaire a cross-national study. *Educational and Psychological Measurement* 66(4):701–716.
- Schwartz, H. A.; Eichstaedt, J. C.; Dziurzynski, L.; Blanco, E.; Kern, M. L.; Kosinski, M.; Stillwell, D.; Seligman, M. E. P.; and Ungar, L. H. 2013. Toward personality insights from language exploration in social media. In *Proceedings of the AAAI Workshop on Analyzing Microtext*.
- Seligman, M. E. P. 2011. *Flourish: A Visionary New Understanding of Happiness and Well-being*. Free Press.
- Smith, A., and Brenner, J. 2012. Twitter use 2012. Technical report, PewResearchCenter.
- Stathopoulou, G.; Powers, M. B.; Berry, A. C.; Smits, J. A.; and Otto, M. W. 2006. Exercise interventions for mental health: a quantitative and qualitative review. *Clinical Psychology: Science and Practice* 13(2):179–193.
- Stiglitz, J.; Sen, A.; and Fitoussi, J.-P. 2009a. The measurement of economic performance and social progress revisited. *Observatoire Français des Conjonctures Economiques (OFCE). The ESPON* 2013.
- Stiglitz, J. E.; Sen, A.; and Fitoussi, J.-P. 2009b. Report by the commission on the measurement of economic performance and social progress. *The Commission on the Measurement of Economic Performance and Social Progress website*. Available online at: <http://www.stiglitz-sen-fitoussi.org>.
- Sumner, C.; Byers, A.; Boochever, R.; and Park, G. 2012. Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. www.onlineprivacyfoundation.org.
- Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* 267–288.
- Vaillant, G. 2008. *Spiritual evolution: A scientific defense of faith*. Harmony.