Understanding and Measuring Psychological Stress Using Social Media

Sharath Chandra Guntuku,[1] **Anneke Buffone,**[1] **Kokil Jaidka**[2]
Johannes C. Eichstaedt,[1] **Lyle H. Ungar**[1]
[1]University of Pennsylvania, [2]Nanyang Technological University
{sharathg@sas, annekeb@sas, jeich@sas, ungar@cis}.upenn.edu,jaidka@ntu.edu.sg

Abstract

A body of literature has demonstrated that users' mental health conditions, such as depression and anxiety, can be predicted from their social media language. There is still a gap in the scientific understanding of how psychological stress is expressed on social media. Stress is one of the primary underlying causes and correlates of chronic physical illnesses and mental health conditions. In this paper, we explore the language of psychological stress with a dataset of 601 social media users, who answered the Perceived Stress Scale questionnaire and also consented to share their Facebook and Twitter data. Firstly, we find that stressed users post about exhaustion, losing control, increased self-focus and physical pain as compared to posts about breakfast, family-time, and travel by users who are not stressed. Secondly, we find that Facebook language is more predictive of stress than Twitter language. Thirdly, we demonstrate how the language based models thus developed can be adapted and be scaled to measure county-level trends. Since county-level language is easily available on Twitter using the Streaming API, we explore multiple domain adaptation algorithms to adapt user-level Facebook models to Twitter language. We find that domain-adapted and scaled social media-based measurements of stress outperform sociodemographic variables (age, gender, race, education, and income), against ground-truth survey-based stress measurements, both at the user- and the county-level in the U.S. Twitter language that scores higher in stress is also predictive of poorer health, less access to facilities and lower socioeconomic status in counties. We conclude with a discussion of the implications of using social media as a new tool for monitoring stress levels of both individuals and counties.

Introduction

Stress is defined as perceived distress caused by an interaction between a person and their environment (Cohen, Miller, and Rabin 2001). While people can handle stress better or worse depending on their general coping skills, experiencing stress too frequently is known to affect well-being and physical and mental health negatively. Stress is seen to be a single pervasive trait influencing health, through a broad range of negative affective states and somatic pathways (McEwen and Stellar 1993). Considering that the symptoms associated with depression and other severe mental health conditions are
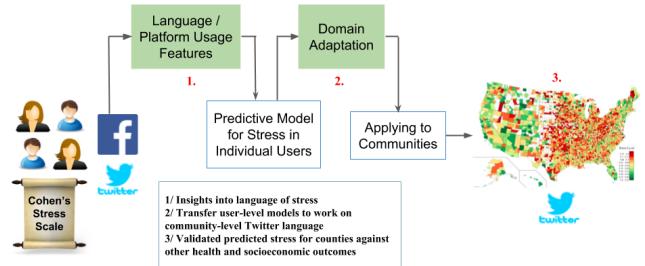
Figure 1: Overview of the approach taken to build language based stress prediction model.

very severe, alleviating psychological stress and promoting a healthy lifestyle is more efficient compared to treating a more acute and chronic condition (Wilkinson 2005).

People are increasingly using social media platforms in order to inform others about their mental states, solicit social support, as well as keep records of their daily activities, preferences, and interests. Notwithstanding the challenge of working with a non-random, non-representative sample of social media users, studies have identified the markers of self-disclosure which concern depression (Guntuku et al. 2017c), schizophrenia (Ernala et al. 2017), ADHD (Guntuku et al. 2017b), alcohol consumption (Liu, Weitzman, and Chunara 2017), and personality (Guntuku et al. 2017a). With respect to stress, the linguistic features of event-related stress have been predicted from social media posts about experiences such as travel and work (Lin et al. 2014); however, these findings cannot be applied to improve the psychological understanding of stress, because people suffering from chronic stress do so irrespective of stressful events. For instance, preparing for an exam is a stressful event, while chronically feeling overwhelmed with responsibilities is trait-related stress. Another research gap is that the previous work has focused on known stressors collected using search keywords (Thelwall 2017). However the labels thus acquired likely have personality confounds (Preotiuc-Pietro et al. 2015), emphasizing the need for using stronger ground truth. Instead, we anticipate that insights into psychological stress could help in (a) designing social-media-based interventions to enable a low-stress lifestyle, and (b) developing a better understanding of regional variations in stress.

Table 1: Items on the Cohen's stress scale: Each question is assessed on a Likert Scale. (-) indicates reverse coded items.

In the last month, how often have you:
- been upset because of something that happened unexpectedly?
- felt that you were unable to control the important things in your life?
- felt nervous and "stressed"?
- felt confident about your ability to handle your personal problems? (-)
- felt that things were going your way? (-)
- found that you could not cope with all the things that you had to do?
- been able to control irritations in your life? (-)
- felt that you were on top of things?(-)
- been angered because of things that were outside of your control?
- felt difficulties were piling up so high that you could not overcome them?

The ubiquitous nature of smart devices and Internet access in almost all parts of the world means that social media is a potentially powerful tool to measure the psychological states and behaviors of people at both micro- (individual) and macro- (county) levels. However, except for a few studies, little has been done to explore how to scale language models to study regions, and no work yet has attempted to do this for stress. Although some studies have analyzed the geographic variation in social media language corresponding to chronic illnesses (Culotta 2014), depression (Bagroy, Kumaraguru, and De Choudhury 2017), well-being (Schwartz et al. 2013a), heart-disease (Eichstaedt et al. 2015), and happiness (Quercia et al. 2012), they are often challenged by (a) a limited understanding of how to scale user-level models to measure counties, and (b) the lack of ground truth about a large population. Recent work has shown the need to use weighting and scaling techniques to transform user-level language estimates from Twitter to county-level estimates (Rieman et al. 2017), in order to avoid the ecological fallacies reported in the studies mentioned above. However, there are challenges associated with transferring predictive models from one social media platform to another, because of differences in self-disclosure on Facebook vs. Twitter (Jaidka et al. 2018). This study contributes with (a) effective off-the-shelf methods for cross-domain adaptation of user-level stress models to measure county-level stress, and (b) validation against ground-truth survey data built on over two million responses.

To summarize, the research gaps in previous work are the lack of language models to predict psychological stress, a limited understanding of scaling and transforming Facebook language models to work on Twitter, and the lack of validation against the region-level ground truth. We show in Figure 1, how our study uses transfer learning to adapt user-level models trained on their Facebook language, to predict county-level stress from the county-level Twitter language. This is necessary because it is easier to train predictive models on the language of a small population of social media users, but it is expensive to survey entire counties for training purposes. Furthermore, county-level social media language is only available for Twitter, where approximately 20% of all public posts are geo-tagged with their location information, and they can be easily mined by using Twitter's Streaming API[1]. On the contrary, Facebook needs user authentication for accessing their posts which is resource intensive to collect

from a large number of people for this research study. Adapting models trained at user-level to predict stress in counties has multiple applications in monitoring health and well-being in counties, especially where survey data is hard to collect. Thus motivated, in this paper we address the following research questions:

RQ1 How does social media language of users who are stressed differ from those who are not?

RQ2 How do Facebook and Twitter language differ in predicting user-level stress? Since county-level language is easily available on Twitter, can off-the-shelf domain adaptation algorithms be used to improve prediction on Twitter language?

RQ3 How do domain-adapted social media-based measurements of stress at the county level correlate with health behaviors and socioeconomic characteristics at the county level?

RQ1: Differential Language Analysis of Stressed Users

Methods

User-level social media data: We deployed a survey on Qualtrics[2] (a platform similar to Amazon Mechanical Turk), comprising several demographic questions (age, gender, race, education, and income) and the Cohen's 10-item Stress scale (Cohen, Kessler, and Gordon 1997) (Table 1), and invited users to share access to their Facebook status updates and/or Twitter usernames. Users received an incentive for their participation, and we obtained their informed consent to access their Facebook and Twitter posts. All users were based in the US. This study received approval from the IRB of our institution.
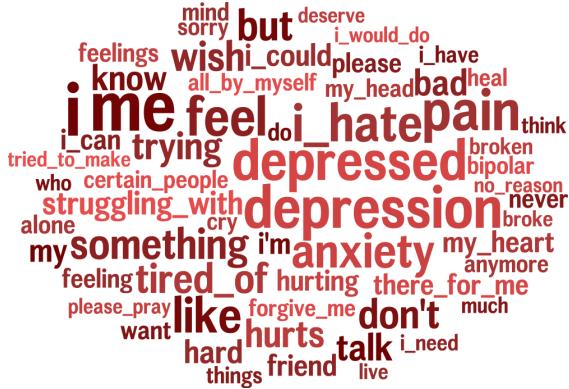
Out of all users who took the survey, 601 users completed the survey and had active accounts with more than 900 words on both Facebook and Twitter. We collected their Facebook posts by using the Facebook Graph API and downloaded their Twitter posts using the Twitter API. Of these 601 users, 265 self-identified as female. The mean age of the sample was 38. The stress scores range from 6 to 39 (mean 30). Each item in the scale is scored on 0-4, with an absolute maximum summing to 40.

Features: We process all our social media posts using the HappierFunTokenizer available with the DLATK package (Schwartz et al. 2017) which is emoticon- and social media-aware. We then represent the language of each user and county as a set of features. In the dictionary-based method, we transform social media language into numerical features representing percentage proportions of lexical categories in an existing dictionary. In the data-driven method, we transform language into numerical features which represent the proportions of word clusters which are statistically similar according to their frequency distributions.

LIWC: We use Linguistic Inquiry and Word Count (LIWC) (Pennebaker, Booth, and Francis 2007), a dictionary

---

[1]https://developer.twitter.com/en

[2]www.qualtrics.com/Survey-Software

Figure 2: Words and phrases associated with a) high-stress users (red) and b) low-stress users (blue). The size of the word indicates the correlation strength and the color indicates frequency (darker is more frequent). Correlations are controlled for age and gender, and are significant at $p < .01$, two-tailed t-test Bonferroni p-correction.

comprising 73 different psycholinguistic categories (e.g., topical categories, emotions, parts-of-speech) to represent the language of each user and county as the normalized frequency distributions.

Topics: In the data-driven method, we represent the language of each user and county as normalized frequency distributions for a set of topics derived using Latent Dirichlet Allocation. These topics are an open-source resource available through the DLATK package (Schwartz et al. 2017) and were trained on a corpus of over 20 million Facebook statuses (Schwartz et al. 2013b).

TensiStrength: To detect direct and indirect expressions of stress or relaxation, a stress lexicon (Thelwall 2017) is used. We obtain stress scores at the sentence level, from each post on Facebook and Twitter and aggregate them to users by calculating mean stress scores.

User engagement We extracted features such as the number of posts made between 12am-6am, mean message length, and number of URLs & hashtags. These features were shown to be predictive of stress by (Lin et al. 2014).

Results

We identify the linguistic characteristics indicative of high and low psychological stress based on the social media language of individual users on Facebook. We conducted a similar analysis on Twitter and present them in the appendix. Since we explore several features simultaneously, we consider coefficients significant if they are less than a Bonferroni-corrected two-tailed p of 0.01 (i.e., when examining 1000 features, in the case of words and phrases, a passing p-value is less than $1x10^{-5}$; when examining 2000 topics p-value is less than $5x10^{-6}$, and when examining 73 LIWC categories p-value is less than $1.3x10^{-4}$).

Ngrams: We extract 1-,2-, and 3-grams from all posts to analyze significant associations between words & phrases and stress. Figure 2 provides a visualization of the Pearson correlations of words and phrases with stress on Facebook. In general, the language of stress has a prominent self-focus. Given that psychologists have found stress to be a state resulting from an individual assessing the demands of the stress-

ful stimulus and determining own coping abilities (Staal 2004), this result is very intuitive. It should be noted that increased self-focus in stressful situations is likely adaptive, but a prolonged self-focus in one's thoughts, especially in the context of negative affect, is known to psychologists as rumination and linked to negative effects for health and well-being (Moberly and Watkins 2008). Hate could signify aggressive or angry affect, which also tends to be maladaptive and could signify the experience of frustration while perceiving resources fall short of perceived demands.

Words and phrases such as "me", "I had", "feel like", "I don't" and "I hate" are significantly correlated with high stress. Furthermore, the language of high stress appears to be marked by expressions of perceived lack of control and expressions of a need state or lack of resources ("struggling with", "tired of", "I need"), as well as negative-angry affect ("hate", "I hate"). Further, high stress language seem to be comorbid with mental health conditions ("depression", "depressed", "anxiety", "bipolar"). It is interesting that language reflects the adverse effects that stress can have on health (Watson and Pennebaker 1989). The language of low stress has prominent positive affect ("excited to"), discussions of meals ("breakfast"), as well as feelings of social inclusion ("joined the"). The language of low stress comprises the discussion of meals, specifically "breakfast", which may indicate relaxation and enjoyment inasmuch as meals are often taken in the company of others, social inclusion, as well as with positive affect ("excited to").

LIWC: LIWC categories from Facebook language that significantly correlate with stress are shown in Table 2. The top correlated categories in Facebook comprise first person singular pronouns ("I"), indicating increased self-focus, which corroborate previous findings (Pennebaker and Lay 2002) that self-references by individuals increase in emotionally vulnerable situations. Users on Facebook are more likely to use adverbs such as "very" and "really" to emphasize their point, and more likely to explicitly use words denoting negative emotions, such as "hurt" and "anger". Mentions of negative emotion confirms our expectations, because stress is an aversive state (Cohen, Kessler, and Gordon 1997). Filler

Table 2: Pearson correlations (top 5) between stress score and LIWC features extracted on Facebook, when controlling for age and gender. All features are significant at $p < .01$, two-tailed t-test and Bonferroni corrected.

LIWC		
Positively correlated	most frequent words	r
1st Person Singular	I	.22
Adverbs	very, really	.17
Negation	no, not, never	.16
Negative Emotion	hurt, sadness, anger	.17
Fillers	I mean, you know	.15
Negatively correlated	most frequent words	r
1st Person Plural	we	-.18
Affiliation	friend, social	-.17
Positive Emotion	love, nice, sweet	-.12

Table 3: Pearson correlations between stress score and Topics extracted from Facebook, when controlling for age and gender. Topic labels are manually created. All features are significant at $p < .01$, two tailed t-test and Bonferroni corrected. The top 5 topics are shown for both positive and negative correlations.

Topics		
Positively correlated	most frequent words used by high-stress people	r
Exhaustion	i'm, tired, hungry, bored, exhausted, freaking, sleepy, soooooooooo, stressed, grumpy	.21
Feeling hurt	i'm, sick, tired, feeling, hearing, tire, fed, bullshit, assuming, hurting, numb	.21
Physical Pain	don't, feel, anymore, beg, begging, honestly, knees, clue, hollow, creeping, sympathy	.20
Feeling sick	feel, sick, crap, feeling, ugh, hate, feels, sucks, crappy, bleh, worse, miserable, sickness, icky, =(	.19
Losing control	i'm, it's, i've, don't, lost, mind, quarter, wouldn't, anymore, control, reason	.19
Negatively correlated	most frequent words used by low-stress people	r
Family and eating	great, lunch, nice, dinner, family, enjoyed, church, wonderful, afternoon, sunday, kids, evening, shopping, meeting, hubby	.13
Travel	kings, delhi, leon, mumbai, rocks, queens, reached, rains, phew, bak, travelling, royal, metro, indians, mahal	-.12

word associations may indicate a lack of self-esteem or feeling down on oneself, and hedging as a result (Abouserie 1994). Negation further suggests a 'lack of' things, mentioned and experienced by those who are also prone to high stress. On the other hand, the first person plural pronouns ("We"), and the LIWC "Affiliation" categories are negatively correlated with stress, which implies that a high-stress individual often depicts themselves as isolated, with a certain disassociation from their social circles. Negative correlations of stress with positive emotion reflects that those individuals reporting lower stress are significantly more likely to express positive emotion.

Topics: Topics which significantly correlate with stress are provided in Table 3. Exhaustion is typical of prolonged stress (McManus, Winder, and Gordon 2002). Feeling hurt, physical pain, and feeling sick are known correlates of stress (Gil et al. 2004). Lack of control also signifies the concept of resources falling short of demands, and possibly reducing their ability to cope with stressors (Gray, Waytz, and Young 2012).

Our findings also reiterate previous research on language use in mental health (De Choudhury et al. 2013), as they describe symptoms of physical pain and sickness, besides expressing a lack of control and negative emotions. On the other hand, the mention of family meals on Facebook has a negative correlation with stress, suggesting a good social support network and taking time to spend with loved ones as well as travel, are ways people relax.

RQ2: Predicting User-Level Psychological Stress using Facebook and Twitter

Methods

We utilize the Facebook and Twitter data from the same users, described in the previous section, to evaluate the performance of supervised models trained on Facebook and Twitter language at predicting users' stress for a held-out set. We stratify our set of users into five folds with a uniform distribution of age and gender traits in each fold. We conduct a cross-validated weighted linear regression for stress, training on

(a) LIWC features and (b) Topic features (c) TensiStrength scores, and (d) engagement features (such as time of posts, number of posts, number of posts between 12am-6am) for users in four folds, and testing on the users in the held out fold.

In the five-fold cross-validation setting, we perform linear regression with several regularization methods such as ridge, elastic-net, LASSO and L2 penalized SVMs and find that elastic-net showed marginally superior performance over the others. Accordingly, we report results only using elastic-net. The performance was measured by calculating Pearson's r over the aggregated predictions from the five folds.

We first evaluate the above features to predict stress *within domain* - i.e., we train and test on the same platform. Engagement features are part of the feature set used in (Lin et al. 2014) for predicting user-level stress. We also examine how the models perform when compared to sociodemographic variables, namely age, gender, race, income, and education.

Then, we evaluate how models trained on Facebook perform at predicting stress from Twitter language in a *cross-domain* setting. Previous studies showed that predictive performance changes in cross-domain applications (Jaidka et al. 2018; Zhong et al. 2017). Therefore, we then attempt to improve the cross-domain prediction performance by applying *domain adaptation*. The motivation for cross-domain and domain adaptation experiments is to build an accurate predictive model on Twitter language to then scale it to county level Twitter language.

Table 4: Within Domain: Stress Prediction Performance (Pearson's r) based on 5-fold cross-validation using different features and models trained and tested on the same domain. Social media language adds to and outperforms sociodemographic variables at predicting stress.

D: Sociodemographic variables		
Feature	Pearson r	
Age, Gender, Race Income & Education	.25	
SM: Social media language		
Feature	Facebook	Twitter
User Engagement (Lin et al. 2014)	.11	.05
TensiStrength (Thelwall 2017)	.17	.11
LIWC 2015	.29	**.22**
Topics	**.31**	.18
Language + Sociodemographic		
D + SM	**.33**	.26

Results for Within Domain Predictions

Table 4 shows the performance of predicting stress using sociodemographic variables and social media language. Within the domain, Facebook does better than Twitter by a slight margin. On Facebook, both LIWC and Topics outperform sociodemographic variables (age, gender, race, income, and education). Topics outperform LIWC on Facebook (r=.305), and LIWC outperforms Topics on Twitter (r=.218). These correlations, which are considered a high correlation in measuring internal traits (Meyer, Finn, and others 2001), show that linguistic features perform reliably well at predicting stress. User engagement features performed rather poorly when compared to linguistic features indicating that trait-prediction is a different task when compared to state-prediction. A similar observation has been made while using user engagement features for predicting other traits (Preotiuc-Pietro et al. 2016). The correlation between stress scores collected from survey and user-aggregated TensiStrength scores was .17 for Facebook and .11 for Twitter. TensiStrength was developed by annotating keyword-selected Twitter posts, which makes them inapplicable for measuring psychological stress. Consequently, we dropped engagement features and TensiStrength from subsequent analysis. We also limited the posting period to the previous one month (consistent with the time period in the Cohen's survey questionnaire) and observed that the performance shows a non-significant drop by 0.02 in r.

Results for Cross-Domain Predictions

Since our goal is to predict stress in counties using Twitter, we examined how models trained on Facebook perform on Twitter (shown in Table 5). Performance drops (by 5% compared to within domain performance) when Facebook models are used to predict stress from the Twitter language. Specifically, topics see a larger drop (by 50%) possibly due to not being able to generalize across platforms. We have seen that expressions of stress in Facebook language vs. Twitter language are significantly different in terms of the vocabulary used. Consequently, a standardized theory-driven dictionary such

as LIWC is stable across both platforms. We also combined both Facebook and Twitter corpora, and the model trained on both platforms together gives a marginal improvement in prediction performance.

Given the marked distinction between expressions of stress in Facebook and Twitter language, we investigate several approaches to improve the predictive performance of models on Twitter. This problem can be viewed as a domain adaptation task, where we are adapting from a *source domain*: users' language on one platform, to a *target domain*: the same users' language on another platform.

Table 5: Cross Domain: Stress Prediction Performance (Pearson's r) based on 5-fold cross-validation. FB: Facebook; Tw: Twitter.

Feature	Trained on FB Tested on Tw	Trained on FB+Tw Tested on Tw
LIWC 2015	.23	**.24**
Topics	.15	.17

Results with Domain adaptation

Most prior work on domain adaptation has focused on the case where some labels are available on both the source and target domains and is usually done by combining (often in some weighted fashion) training data sets or, less commonly, trained models from the source and target domains. A simple and effective supervised method was proposed by Daumé (Daumé III, Kumar, and Saha 2010) which applied a supervised heuristic mapping from labeled data the source and target domains, to a higher dimensional feature space, which are used to train standard classifiers or linear regression models. This approach has demonstrated the best performance in several comparative evaluations conducted for image-, text- and sentiment-classification (Pan and Yang 2010).

We test two standard domain adaptation techniques for improving the cross-platform performance of predictive models: one supervised approach; Easy Adapt (Daumé III, Kumar, and Saha 2010), which uses labeled data from both the source and the target, and one unsupervised approach: Transfer Component Analysis – TCA (Pan et al. 2011), which requires no labels on the target domain. While there are several candidate algorithms within both supervised and unsupervised approaches, we chose Easy Adapt and TCA for their simplicity in application.

The standard notation used throughout this section is that X_S refers to labeled observations in the source domain and X_T refers to the test set in the target domain.

EasyAdapt : We define our problem according to the implementation described by Daumé (Daumé III, Kumar, and Saha 2010). Let $X1$ denote the original feature space for the user in the source domain, $X1 = \mathbb{R}^F$. We construct an augmented feature space $\widetilde{X1} = \mathbb{R}^{3F}$, by creating an platform-specific, and user-specific version of each feature in X. For this, we define $\Phi_s : X \longrightarrow \widetilde{X}$ to transform feature vectors corresponding to the platform-specific and user-specific

feature spaces. The mappings are defined by the following equation:

$$\Phi^s(x) = \langle x, x, \mathbf{0} \rangle, \Phi^t(x) = \langle x, \mathbf{0}, x \rangle \qquad (1)$$

Here, $\Phi^s(x)$ is the feature vector for the source domain, $\Phi^t(x)$ is the corresponding vector for the target domain. $\mathbf{0} = \langle 0, 0, ..., 0 \rangle \in \mathbb{R}^F$ is the zero vector.

We thus augment the original feature space with the labeled data points (for the same set of individuals) from the target domain, excluding the held out sample for testing. Next, we train a regression model between this augmented feature space and stress scores of the users. We similarly transform the feature space of the held out sample before prediction.

Transfer Component Analysis (TCA) TCA exploits the Maximum Mean Discrepancy Embedding (MMDE) for comparing the distributions between the source and target domain, based on the Reproducing Kernel Hilbert Space (RKHS). The empirical estimate of the distance between D_S and D_T, $Dist(D_S, D_T)$, is

$$|| \frac{1}{n_1} \sum_{i=1}^{n_1} s_i - \frac{1}{n_2} \sum_{i=1}^{n_2} u_i ||_H \qquad (2)$$

where u and s are individual observations in D_S and D_U, H is a universal RKHS and $\phi : X \longrightarrow$ H.

After applying domain adaptation (results in Table 6), performance increases by 16% when compared to using Facebook models alone to predict stress from Twitter language without any domain adaptation. TCA is seen to outperform EasyAdapt.

Table 6: Domain Adaptation: Stress Prediction Performance (Pearson's r) based on 5-fold cross-validation, trained on Facebook and tested on Twitter.

Feature	No Domain Adaptation	EasyAdapt	Transfer Component Analysis
LIWC 2015	.23	.23	**.27**
Topics	.15	.16	.18

RQ3 Evaluating County Level Language-Predicted Stress

Methods

Data The language of US counties is obtained from Twitter. It comprises of the normalized frequency distributions of words, averaged by the number of individuals who spoke those words, extracted from a set of geo-located tweets (Schwartz et al. 2013a). The authors constructed this dataset by collecting a 10% sample of Twitter posts between the years 2009-15 from the TrendMiner project (Preotiuc-Pietro et al. 2012). Tweets were geo-mapped to US counties using latitude/longitude coordinates and the self-reported location field when available, following the approach described in (Schwartz et al. 2013a). In this manner, 20% of the tweets could be successfully mapped, resulting in over 1.7 billion geo-located tweets. A threshold of 100,000 words per county

is used in order to ensure that any word-to-outcome correlations observed were stable, resulting in a dataset of 2710 counties. On an average, each county had 8,892,568 words.

County Health outcomes: We use two sets of county health outcomes in our work: 1. Gallup-Sharecare Well-Being data and 2. County health statistics for the United States.

From the Gallup-Sharecare Well-Being Index, we use the stress outcome to validate the stress predictions made by our language model. Gallup data is collected as a part of 1,000 telephone interviews conducted every day across the US (Huppert and So 2010). The questions on the interview range from topics such as health behaviors, work environment to social and county factors, and financial security. We specifically used the stress fields from the Gallup data aggregated to counties.

From US County Health Rankings and Roadmaps portal[3], we obtained county socioeconomic characteristics and health behaviors which provides access to county-level health factors from a wide range of sources, including Behavioral Risk Factor Surveillance System, American counties Survey, and the National Center for Health Statistics.

Results

After analyzing the performance of different models at predicting stress in the cross-domain setting, we use the domain-adapted model trained on LIWC features to predict stress using Twitter language from US counties (shown in Figure 3). We validate our language-predicted stress with county-level stress reported by Gallup, and further with county-level estimates of health behaviors and socioeconomic characteristics.
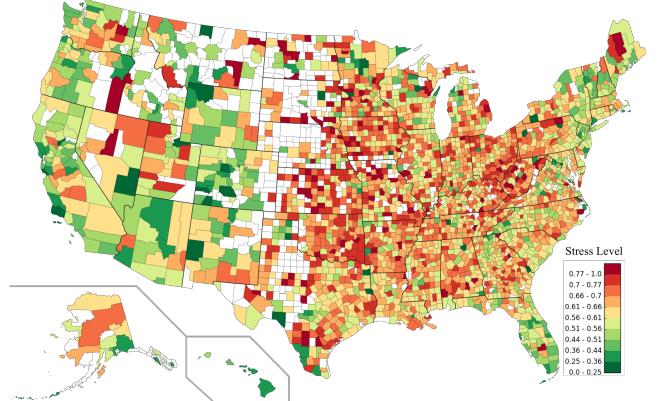

Figure 3: Map showing stress predicted from Twitter language from US counties. Counties which had less than 100k tweets are shown white.

First, we use the Gallup-Sharecare Well-Being data, with county-aggregated stress labels (termed as 'Gallup stress' hereon), as a means of validating our language-predicted stress. We compare Gallup stress and language-predicted stress (referred to as 'Twitter stress' hereon). We aim to train the best performing user-level model, as this directly impacts the county-level prediction accuracy. We experiment with models trained on individual platforms (best $r = 0.26$), and

---

[3]www.countyhealthrankings.org/

find a significant difference in predicting Gallup stress after domain adaptation (r = 0.34). These correlations are considered a high correlation in psychology, especially when measuring outcomes like stress (Meyer, Finn, and others 2001). Part of this improvement is due to the amount of information that Facebook contains over Twitter (Table 4). However, Facebook alone is insufficient as the Twitter language is substantially different from Facebook (Tables 5 and 6).

Predictive performance of sociodemographic variables (age, gender, income and education) and Twitter language are r=.24 and r=.34 respectivelt. Twitter language outperforms sociodemographic variables at predicting stress in counties.

Having validated the language-estimated stress in counties, we examine the correlation of stress with socioeconomic characteristics and health behaviors in US counties. The results are shown in Table 7.

Health Behaviors Counties with higher stress have higher mortality, a higher percentage of physically inactive people, a higher percentage of poor mental health days, and a higher percentage of smokers, teen pregnancy, and higher rates of drug poisoning. These are all factors which have known correlation and causality structures as reported by (Stults-Kolehmainen and Sinha 2014; Steptoe et al. 1996; Watson and Pennebaker 1989). Individuals living in high-stress counties appear to suffer from poor health, affected both by aspects in their individual lives as well as by the nature of their counties. Using social media language as a real-time proxy to assess a county's stress, and consequently, health, could help design interventions and policies to alleviate stressors. Language has the specific advantage of getting at the causes/stressors that counties are experiencing, and this could greatly complement and also inform traditional survey mechanisms (such as Gallup).

Socioeconomic Characteristics Counties which are rich and educated express less stress (Karlamangla, Singer, and Seeman 2006). This is observed by the low correlation between language-predicted stress and Median Household Income (r=-.271) and Education level, % people who attended some college, (r=.345) in counties. Furthermore, counties where people have access to exercise facilities also have low stress (Salmon 2001). Our language-based findings of stress and income are corroborated by an analysis of the Gallup data reported by (De Neve and Ward 2017). High-stress counties also have higher avoidable hospitalization rates and have less access to providers for mental and physical health. These associations suggest that individuals in high-stress counties also face other economic and infrastructural challenges.

Discussion

Our work provides some of the first insights into the linguistic manifestation of psychological stress on social media, while also showing the feasibility of applying models trained at the user-level to predict stress in counties with validity. Language associated with stress is indicative of exhaustion (McManus, Winder, and Gordon 2002), comorbid-

Table 7: Correlation between county level health measures and stress. All correlations are significant at $p < 0.01$. Language-based stress predictions predict health behaviors and clinical care facilities above and beyond socio-demographic controls

County-level Outcomes	Language Correlation	
	With No Controls	With SES controls
Health Behaviors		
MV Mortality Rate	.40	.11
% Physically Inactive	.44	.31
% Smokers	.29	.30
% Poor mental health days	.14	.13
Teen Birth Rate	.23	-
Drug Poisoning Mortality Rate	.13	-
Clinical Care		
Preventable Hosp. Rate	.41	.36
Mental health providers rate	-.39	-.19
Primary Care Physician Rate	-.43	-.33
% Could not see doctor due to cost	.17	-
Socioeconomic Environment		
Median Household Income (ln)	-.27	-
Education (% College graduates)	-.34	-
% With Access to Exercise Facilities	-.45	-.264

ity with mental health conditions (Watson and Pennebaker 1989), pronounced self-focus (Pennebaker and Lay 2002) and feelings of hurt, physical pain, and being sick, which also known correlates of stress (Gil et al. 2004). Our findings also reiterate previous research on language use in mental health (De Choudhury et al. 2013). An important insight from our work was that the predictive utility of different platforms varies, as has been seen in prior research looking at differences in self-disclosure (Jaidka et al. 2018; Zhong et al. 2017). This motivates the need for transfer learning when the stress model trained on Facebook language needs to be applied to Twitter language. This observation has implications for other researchers interested in training and applying machine learning models across domains.

On applying the adapted stress model to the language of counties, we obtain face-valid results about the relationships between stress and ground-truth statistics about county health. Our findings clarify the association of stress with socioeconomic factors (Baum, Garofalo, and Yali 1999), suggesting that it is the socially and economically deprived counties which are likely to face higher stress. Urban counties, with pressures of work and a fast-paced routine, are less likely to experience debilitating, psychological stress (Lederbogen et al. 2011). This is an important way to differentiate the stressful events in an urban lifestyle with the trait-based stress experienced in deprived or rural counties.

Limitations and Future Work

While there are several advantages and merits of using social media to measure stress, one of the apparent shortcomings of linguistic analysis is that they do not offer causal insights. Our

methods anticipate that the language of counties is indicative of stress. It is possible that people might stop using social media while they are stressed, which could impact the performance of our user-level model (Corrigan et al. 2001). However, we did not find any correlation (positive or negative) between the raw number of posts and psychological stress. To address such scenarios, monitoring multiple data sources such as phone-based sensors can be useful in promptly reaching out to the afflicted individuals (Singh and Agarwal 2016). Further, it would be interesting to disentangle the language indicative of short-term stressors versus long-term chronic stress, potentially by collecting data around recent stressful events (divorce, moving house, losing job etc.).

Some studies have shown preliminary findings that connect social media usage to poorer mental health, indicating that undesirable usage of social media might cause depression among its users (O'Keeffe, Clarke-Pearson, and others 2011). While we can see that social media can uncover several behavioral and psychological insights into the stressful lifestyles of people, further research is warranted to precisely determine the causality pathways.

While one way of representing counties is by their geographic distribution (as in the case of US counties), another way could be by their residents' occupation ('work' counties), and students ('college' counties). Monitoring stress levels in such counties would especially be useful in devising programs or initiatives to encourage a low-stress lifestyle amongst wherever individuals are particularly vulnerable to stress. While social media (and Twitter in particular) are not representative of the real-life users, several insights have been obtained using public Twitter data in public health (Yang et al. 2018). Further, the underlying predictive model can be applied to several online forums. For example, it could be applied to colleges across the U.S. to understand the specific chronic stressors in different college campuses (extending the recent work (Saha and De Choudhury 2017; Bagroy, Kumaraguru, and De Choudhury 2017)). It is especially promising that such measurements can be effective while being unobtrusive. Monitoring stress using social media also offers the advantage of discovering stress associations that are otherwise not easy to get by traditional surveys (Holdeman 2009). Moreover, such a tool could monitor stress in real time, potentially detecting unforeseen events before they happen (Reason 1995).

On the technical front, while our application of domain adaptation bridged the gap between two social media platforms – Facebook and Twitter, we note that further work is warranted to address the challenges associated with spatially correlated terms introduced when aggregating tweets to counties. The study by (Rieman et al. 2017) has explored such problems and their implications for county-level predictions from language. Also, several works that focus on socio-geographical studies have used advanced autoregressive modeling techniques which could help in developing more robust models, even for counties with missing data based on their spatial neighborhood association with other counties that have data points. Further, it would be interesting to examine stress from multiple modalities (text, images and sensor data), which can potentially complementary insights

(Sharma and Gedeon 2012; Guntuku, Roy, and Weisi 2015; Samani et al. 2018), particularly transfering learning across modalities (Guntuku et al. 2019).

Privacy and Ethical Considerations

Prediction performances obtained in this study indicate that psychological stress can be inferred with some accuracy from public (Twitter) or semi-public (Facebook) social media data. While the purpose of delivering social support and mental health services motivates this effort, these algorithms also raise some important privacy and ethical questions.

From the perspective of privacy concerns, organizations with vested interests (e.g. insurance companies) may be motivated to infer this information automatically. As being chronically stressed can lead to stigma at multiple levels, data protection and ownership frameworks are necessary to make sure the data is not used against the users interest (McKee 2013). Few users realize the amount of health-related information that can be inferred from their digital traces; consequently, transparency about the indicators derived and the purposes behind such inference should be part of ethical and policy discourse.

Conclusion

Using a sample of users who have accounts on both Twitter and Facebook, we uncovered insight into the language of stress in individuals. Our results also showed that standardized linguistic dictionaries, such as LIWC, outperform engagement attributes such as user posting behavior, in predicting stress. Our results complement psychological survey data by deepening our understanding of the environment that may contribute to both individual and county-level well-being. Language estimates can be used to monitor stress in different social settings, to devise initiatives encouraging a low-stress lifestyle. Our model will help measure the impact of such interventions in real time. Further, language has the specific advantage of getting at the causes/stressors that counties are experiencing thereby enabling personalized interventions.

The findings shed light on the nuanced relationship of socioeconomic status with psychological stress and suggest insights for other psychological factors, such as locus of control, the degree to which people feel like being manipulated by their environment vs. perceive a sufficient level of self-control and in the counties where they live and work. We believe that the potential for this technology to be used in both micro and macro scales is tremendous. Our computational techniques, while useful to monitor stress in counties, can also be used to give real-time feedback to individuals in those counties. Such feedback about individuals' social media usage along specific dimensions, can help in designing simple but effective technology-assisted interventions in cultivating mindfulness and stress control (Miller, Fletcher, and Kabat-Zinn 1995).

References

[Abouserie 1994] Abouserie, R. 1994. Sources and levels of stress in relation to locus of control and self esteem in

university students. *Educational psychology* 14(3):323–330.

[Bagroy, Kumaraguru, and De Choudhury 2017] Bagroy, S.; Kumaraguru, P.; and De Choudhury, M. 2017. A social media based index of mental well-being in college campuses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1634–1646. ACM.

[Baum, Garofalo, and Yali 1999] Baum, A.; Garofalo, J.; and Yali, A. 1999. Socioeconomic status and chronic stress: does stress account for ses effects on health? *Annals of the New York Academy of Sciences* 896(1):131–144.

[Cohen, Kessler, and Gordon 1997] Cohen, S.; Kessler, R. C.; and Gordon, L. U. 1997. *Measuring stress: A guide for health and social scientists*. Oxford University Press on Demand.

[Cohen, Miller, and Rabin 2001] Cohen, S.; Miller, G. E.; and Rabin, B. S. 2001. Psychological stress and antibody response to immunization: a critical review of the human literature. *Psychosomatic medicine*.

[Corrigan et al. 2001] Corrigan, P. W.; Green, A.; Lundin, R.; Kubiak, M. A.; and Penn, D. L. 2001. Familiarity with and social distance from people who have serious mental illness. *Psychiatric services* 52(7):953–958.

[Culotta 2014] Culotta, A. 2014. Estimating county health statistics with twitter. In *CHI*. ACM.

[Daumé III, Kumar, and Saha 2010] Daumé III, H.; Kumar, A.; and Saha, A. 2010. Frustratingly easy semi-supervised domain adaptation. In *Proceedings of the 2010 Workshop on Domain Adaptation for Natural Language Processing*, 53–59. Association for Computational Linguistics.

[De Choudhury et al. 2013] De Choudhury, M.; Gamon, M.; Counts, S.; and Horvitz, E. 2013. Predicting depression via social media. In *ICWSM*, 2.

[De Neve and Ward 2017] De Neve, J., and Ward, G. 2017. Does work make you happy? evidence from the world happiness report. *Harvard Business Review*.

[Eichstaedt et al. 2015] Eichstaedt, J. C.; Schwartz, H. A.; Kern, M. L.; Park, G.; Labarthe, D. R.; Merchant, R. M.; Jha, S.; Agrawal, M.; Dziurzynski, L. A.; Sap, M.; Weeg, C.; Larson, E. E.; Ungar, L. H.; and Seligman, M. E. 2015. Psychological Language on Twitter Predicts County-level Heart Disease Mortality. *Psychological Science* 26:159–169.

[Ernala et al. 2017] Ernala, S. K.; RIZVI, A. F.; BIRNBAUM, M. L.; KANE, J. M.; and DE CHOUDHURY, M. 2017. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. *Proc ACM Hum-Comput Interact* 1(1):43.

[Felson, Scheff, and Retzinger 1993] Felson, R. J.; Scheff, T. J.; and Retzinger, S. M. 1993. Review essay: Shame, anger, and aggression.

[Gil et al. 2004] Gil, K. M.; Carson, J. W.; Porter, L. S.; Scipio, C.; Bediako, S. M.; and Orringer, E. 2004. Daily mood and stress predict pain, health care use, and work activity in african american adults with sickle-cell disease. *Health Psychology* 23(3):267.

[Gray, Waytz, and Young 2012] Gray, K.; Waytz, A.; and Young, L. 2012. The moral dyad: A fundamental template

unifying moral judgment. *Psychological Inquiry* 23(2):206–215.

[Guntuku et al. 2017a] Guntuku, S. C.; Lin, W.; Carpenter, J.; Ng, W. K.; Ungar, L. H.; and Preoţiuc-Pietro, D. 2017a. Studying personality through the content of posted and liked images on twitter. In *Proceedings of the 2017 ACM on web science conference*, 223–227. ACM.

[Guntuku et al. 2017b] Guntuku, S. C.; Ramsay, J. R.; Merchant, R. M.; and Ungar, L. H. 2017b. Language of adhd in adults on social media. *Journal of attention disorders* 1087054717738083.

[Guntuku et al. 2017c] Guntuku, S. C.; Yaden, D. B.; Kern, M. L.; Ungar, L. H.; and Eichstaedt, J. C. 2017c. Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences* 18:43–49.

[Guntuku et al. 2019] Guntuku, S. C.; Preoţiuc-Pietro, D.; Eichstaedt, J. C.; and Ungar, L. 2019. What twitter profile and posted images reveal about depression and anxiety. In *Thirteenth International AAAI Conference on Web and Social Media*.

[Guntuku, Roy, and Weisi 2015] Guntuku, S. C.; Roy, S.; and Weisi, L. 2015. Evaluating visual and textual features for predicting user likes. In *2015 IEEE International Conference on Multimedia and Expo (ICME)*, 1–6. IEEE.

[Holdeman 2009] Holdeman, T. C. 2009. Invisible wounds of war: Psychological and cognitive injuries, their consequences, and services to assist recovery. *Psychiatric Services* 60(2):273–273.

[Huppert and So 2010] Huppert, F. A., and So, T. T. 2010. Flourishing across europe: Application of a new conceptual framework for defining well-being. *Gallup*.

[Jaidka et al. 2018] Jaidka, K.; Guntuku, S. C.; Buffone, A.; Schwartz, H. A.; and Ungar, L. H. 2018. Facebook vs. twitter: Cross-platform differences in self-disclosure and trait prediction. In *ICWSM*.

[Karlamangla, Singer, and Seeman 2006] Karlamangla, A. S.; Singer, B. H.; and Seeman, T. E. 2006. Reduction in allostatic load in older adults is associated with lower all-cause mortality risk: Macarthur studies of successful aging. *Psychosomatic medicine* 68(3):500–507.

[Lederbogen et al. 2011] Lederbogen, F.; Kirsch, P.; Haddad, L.; Streit, F.; Tost, H.; Schuch, P.; Wüst, S.; Pruessner, J. C.; Rietschel, M.; Deuschle, M.; et al. 2011. City living and urban upbringing affect neural social stress processing in humans. *Nature* 474(7352):498.

[Lin et al. 2014] Lin, H.; Jia, J.; Guo, Q.; Xue, Y.; Li, Q.; Huang, J.; Cai, L.; and Feng, L. 2014. User-level psychological stress detection from social media using deep neural network. In *ACM MM*.

[Linn et al. 1985] Linn, L. S.; Yager, J.; Cope, D.; and Leake, B. 1985. Health status, job satisfaction, job stress, and life satisfaction among academic and clinical faculty. *Jama* 254(19):2775–2782.

[Liu, Weitzman, and Chunara 2017] Liu, J.; Weitzman, E. R.; and Chunara, R. 2017. Assessing behavioral stages from social media data. In *Proceedings of the 20th ACM Conference*

Companion on Computer Supported Cooperative work & social computing (CSCW), volume 2017, 1320. NIH Public Access.

[McEwen and Stellar 1993] McEwen, B. S., and Stellar, E. 1993. Stress and the individual: mechanisms leading to disease. *Archives of internal medicine* 153(18):2093–2101.

[McKee 2013] McKee, R. 2013. Ethical issues in using social media for health and health care research. *Health Policy* 110(2):298–301.

[McManus, Winder, and Gordon 2002] McManus, I.; Winder, B.; and Gordon, D. 2002. The causal links between stress and burnout in a longitudinal study of uk doctors. *The Lancet* 359(9323):2089–2090.

[Meyer, Finn, and others 2001] Meyer, G. J.; Finn, S. E.; et al. 2001. Psychological Testing and Psychological Assessment: A Review of Evidence and Issues. *American Psychologist*.

[Miller, Fletcher, and Kabat-Zinn 1995] Miller, J. J.; Fletcher, K.; and Kabat-Zinn, J. 1995. Three-year follow-up and clinical implications of a mindfulness meditation-based stress reduction intervention in the treatment of anxiety disorders. *General hospital psychiatry* 17(3):192–200.

[Moberly and Watkins 2008] Moberly, N. J., and Watkins, E. R. 2008. Ruminative self-focus, negative life events, and negative affect. *Behaviour research and therapy* 46(9):1034–1039.

[O'Keeffe, Clarke-Pearson, and others 2011] O'Keeffe, G. S.; Clarke-Pearson, K.; et al. 2011. The impact of social media on children, adolescents, and families. *Pediatrics* 127(4):800–804.

[Pan and Yang 2010] Pan, S. J., and Yang, Q. 2010. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22(10):1345–1359.

[Pan et al. 2011] Pan, S. J.; Tsang, I. W.; Kwok, J. T.; and Yang, Q. 2011. Domain adaptation via transfer component analysis. *IEEE TNN*.

[Pennebaker and Lay 2002] Pennebaker, J. W., and Lay, T. C. 2002. Language use and personality during crises: Analyses of mayor rudolph giuliani's press conferences. *Journal of Research in Personality*.

[Pennebaker, Booth, and Francis 2007] Pennebaker, J. W.; Booth, R. J.; and Francis, M. E. 2007. Linguistic inquiry and word count: Liwc [computer software]. *Austin, TX: liwc. net*.

[Preotiuc-Pietro et al. 2012] Preotiuc-Pietro, D.; Samangooei, S.; Cohn, T.; Gibbins, N.; and Niranjan, M. 2012. Trendminer: An architecture for real time analysis of social media text. In *Sixth International AAAI Conference on Weblogs and Social Media*.

[Preotiuc-Pietro et al. 2015] Preotiuc-Pietro, D.; Eichstaedt, J.; Park, G.; Sap, M.; Smith, L.; Tobolsky, V.; Schwartz, H. A.; and Ungar, L. 2015. The role of personality, age and gender in tweeting about mental illnesses. In *NAACL HLT*, volume 2015, 21.

[Preotiuc-Pietro et al. 2016] Preotiuc-Pietro, D.; Carpenter, J.; Giorgi, S.; and Ungar, L. 2016. Studying the dark triad of personality through twitter behavior. In *CIKM*, 761–770. ACM.

[Quercia et al. 2012] Quercia, D.; Ellis, J.; Capra, L.; and Crowcroft, J. 2012. Tracking gross community happiness from tweets. In *Proceedings of the 15th ACM Conference Companion on Computer Supported Cooperative work & social computing (CSCW)*. ACM.

[Reason 1995] Reason, J. 1995. Understanding adverse events: human factors. *BMJ Quality & Safety* 4(2):80–89.

[Rieman et al. 2017] Rieman, D.; Jaidka, K.; Schwartz, H. A.; and Ungar, L. 2017. Domain adaptation from user-level facebook models to county-level twitter predictions. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, 764–773.

[Rouhizadeh et al. 2016] Rouhizadeh, M.; Ungar, L.; Buffone, A.; and Schwartz, H. A. 2016. Using syntactic and semantic context to explore psychodemographic differences in self-reference. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2054–2059.

[Saha and De Choudhury 2017] Saha, K., and De Choudhury, M. 2017. Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction* (CSCW).

[Salmon 2001] Salmon, P. 2001. Effects of physical exercise on anxiety, depression, and sensitivity to stress: a unifying theory. *Clinical psychology review* 21(1):33–61.

[Samani et al. 2018] Samani, Z. R.; Guntuku, S. C.; Moghaddam, M. E.; Preoţiuc-Pietro, D.; and Ungar, L. H. 2018. Cross-platform and cross-interaction study of user personality based on images on twitter and flickr. *PloS one* 13(7):e0198660.

[Schwartz et al. 2013a] Schwartz, H. A.; Eichstaedt, J. C.; Kern, M. L.; Dziurzynski, L.; Lucas, R. E.; Agrawal, M.; Park, G. J.; Lakshmikanth, S. K.; Jha, S.; Seligman, M. E. P.; and Ungar, L. H. 2013a. Characterizing Geographic Variation in Well-being using Tweets. ICWSM.

[Schwartz et al. 2013b] Schwartz, H. A.; Eichstaedt, J. C.; Kern, M. L.; Dziurzynski, L.; Ramones, S. M.; Agrawal, M.; Shah, A.; Kosinski, M.; Stillwell, D.; and Seligman, M. 2013b. Personality, Gender, and Age in the Language of Social Media: The Open-vocabulary Approach. *PloS ONE*.

[Schwartz et al. 2017] Schwartz, H. A.; Giorgi, S.; Sap, M.; Crutchley, P.; Ungar, L.; and Eichstaedt, J. 2017. Dlatk: Differential language analysis toolkit. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 55–60.

[Sharma and Gedeon 2012] Sharma, N., and Gedeon, T. 2012. Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer methods and programs in biomedicine* 108(3):1287–1301.

[Singh and Agarwal 2016] Singh, V. K., and Agarwal, R. R. 2016. Cooperative phoneotypes: exploring phone-based behavioral markers of cooperation. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 646–657. ACM.

[Staal 2004] Staal, M. A. 2004. Stress, cognition, and human performance: A literature review and conceptual framework.

[Steptoe et al. 1996] Steptoe, A.; Wardle, J.; Pollard, T. M.; Canaan, L.; and Davies, G. J. 1996. Stress, social support and health-related behavior: a study of smoking, alcohol consumption and physical exercise. *Journal of psychosomatic research* 41(2):171–180.

[Stults-Kolehmainen and Sinha 2014] Stults-Kolehmainen, M. A., and Sinha, R. 2014. The effects of stress on physical activity and exercise. *Sports medicine* 44(1):81–121.

[Thelwall 2017] Thelwall, M. 2017. Tensistrength: Stress and relaxation magnitude detection for social media texts. *IPM*.

[Watson and Pennebaker 1989] Watson, D., and Pennebaker, J. W. 1989. Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychological review* 96(2):234.

[Wilkinson 2005] Wilkinson, R. G. 2005. *The impact of inequality: How to make sick societies healthier*. The New Press.

[Yang et al. 2018] Yang, Q.; Tufts, C.; Ungar, L.; Guntuku, S.; and Merchant, R. 2018. To retweet or not to retweet: Understanding what features of cardiovascular tweets influence their retransmission. *Journal of health communication* 1–10.

[Zhong et al. 2017] Zhong, C.; Chan, H.-w.; Karamshu, D.; Lee, D.; and Sastry, N. 2017. Wearing many (social) hats: How different are your different social network personae? ICWSM.

Appendix

Here, we describe the experiments detailing the language differences between Facebook and Twitter and their associations with stress.

Usage Statistics: In terms of pure usage statistics, we find that the our 601 participants had an average of 1238 Facebook and 1649 Twitter posts. The mean number of words they wrote per post was 17.1 for Twitter and 23.8 for Facebook. This difference is expected given the 140-character limit imposed on Twitter posts. However, the *median* number of words per post – 17 for Twitter and 18 for Facebook – is much closer, suggesting that our participants were more varied in their posting behavior, when it came to Facebook.

N Grams: Next, we consider platform differences in expressing stress, at the word- and phrase-level. Figure 4 provides a 2D visualization of the Spearman correlations of individual words with stress on the X-axis, and with social media platform on the Y-axis. The words along the X-axis are highly correlated with stress, but less preferential towards individual social media platforms. The words at the Y-axis poles are highly correlated with platform usage but less predictive of stress. For the purpose of this plot, we have discarded words with non-significant correlations as well as $\rho <= 0.05$. Thus, the words along the diagonals are the ones which best distinguish Twitter usage from Facebook usage for the purpose of expressing stress.

On Facebook, words and phrases such as "me", "i had", "feel like", "i don't" and "i hate" are significantly correlated with high stress, while words and phrases such as "lunch","a great year" and "to see" are significantly correlated with low stress. On the other hand, on Twitter, words and phrases which indicate low stress are "excited to", "! check", "fresh", "win" and "experience".

The lack of words in the bottom right quadrant suggests that that language of high stress on Twitter is a subset of the language of high stress on Facebook. These findings motivate our proposal to adapt stress models trained on Facebook language, a more inclusive corpus, to Twitter language, for the purpose of county-level stress prediction.

LIWC: From Table 8, we observe that Twitter has some similarities with Facebook - such as the use of adverbs, and the negative correlation with the use of "Positive Emotion" words. Unexpectedly, on Twitter, the category most strongly correlated with stress is "Comparisons" and not first person singular pronouns. We anticipate that this reflects the tendency of Twitter users to drop self-referencing pronouns, as has been identified in previous work (Rouhizadeh et al. 2016). Comparisons and differentiation could indicate lack of satisfaction with status quo, indicating possibly pessimism or expressing perceived resources to fall short of demands, or simply lower perceived satisfaction with life of those prone to stress (Linn et al. 1985). Insecurities of those higher in stress may be what is indicated by correlations with nonfluencies (Felson, Scheff, and Retzinger 1993). A "focus on the past" tends to be maladaptive and may indicate that those high in stress might find themselves stuck in what happened and in ruminative, self-deprecating thoughts. This is reinforced by the negatively correlated category "Power", which suggests that the individuals under stress express a lack of control and superiority over their environment (Abouserie 1994).

Topics: Twitter topics which are positively correlated with stress (Table 9) comprise of words and emojis depicting negative emotions, possible sarcasm and feelings of awkwardness. We did not find any topics which are negatively correlated with stress, on Twitter when controlled for age and gender.

TensiStrength: The Pearson correlation between the mean stress values for Facebook and Twitter, was 0.28 and significant at $p < 0.01$. This difference is expected because the TensiStrength lexicon was developed on Twitter posts, and as such it might not generalize well to Facebook language. The finding does suggest that stress is expressed very differently on Twitter versus Facebook, and motivates our decision to adapt Facebook stress models to Twitter.
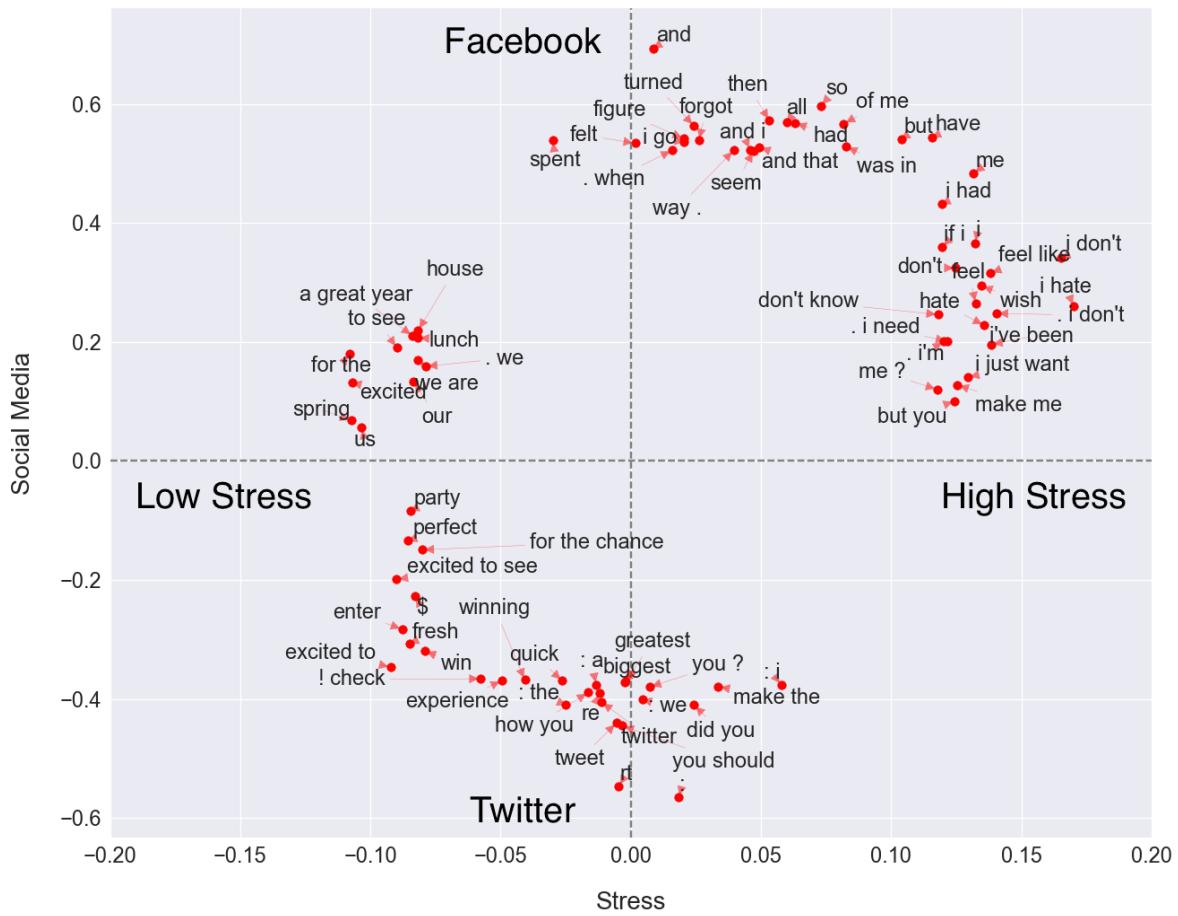
Figure 4: Correlations of words/phrases with Social Media platform (Facebook on the positive y-axis and Twitter on the negative y-axis) and Stress Scale. Correlations are significant, Bonferroni corrected at $p < 0.01$

LIWC - Twitter		
Positively correlated	most frequent words	r
Comparisons	greater, best, after	.164
Adverbs	very, really	.144
Nonfluencies	er, hm, umm	.136
Differentiation	hasn't, but, else	.134
Focused on Past	ago, did, talked	.134
Negatively correlated	most frequent words	r
Power	superior	-.159
Positive Emotion	love, nice, sweet	-.107

Table 8: Pearson correlations (maximum 5) between the stress score and LIWC features extracted on Twitter, when controlling for age and gender. All features are significant at $p < .01$, two tailed t-test.

Topics – Twitter		
Positively correlated	most frequent words used by high-stress people	r
Hatred	hate, stupid, $-\_$, $>$. $<$, gah, ugh, friggin, effin, sigh, grr, $>$:,wtf, annoying, urgh, dammit	.165
Sarcasm	funny, it's, thinks, people, isn't, finds, hilarious, ironic, entertaining, nicer, unfortunate	.159
Feeling awkward	feels, weird, kinda, weird, bit, strange, odd, hmm, suddenly, sort, awkward, dunno	.150

Table 9: Pearson correlations between the stress score and Topics extracted from Twitter posts, when controlling for age and gender. Topic labels are manually created. All features are significant at $p < .01$, two tailed t-test. No negative correlations were found for Twitter posts