A Framework for Understanding the Relationship between Social Media Discourse and Mental Health

SANJANA MENDU, University of Virginia, Charlottesville, VA USA
ANNA BAGLIONE, University of Virginia, Charlottesville, VA USA
SONIA BAEE, University of Virginia, Charlottesville, VA USA
CONGYU WU, University of Texas at Austin, Austin, TX USA
BRANDON NG, University of Richmond, Richmond, VA USA
ADI SHAKED, University of Virginia, Charlottesville, VA USA
GERALD CLORE, University of Virginia, Charlottesville, VA USA
MEHDI BOUKHECHBA, University of Virginia, Charlottesville, VA USA
LAURA BARNES, University of Virginia, Charlottesville, VA USA

Over 35% of the world's population uses social media. Platforms like Facebook, Twitter, and Instagram have radically influenced the way individuals interact and communicate. These platforms facilitate both public and private communication with strangers and friends alike, providing rich insight into an individual's personality, health, and wellbeing. To date, many researchers have employed a variety of methods for extracting mental health-centric features from digital text communication (DTC) data, including natural language processing, social network analysis, and extraction of temporal discourse patterns. However, none have explored a hierarchical framework for extracting features from private messages with the goal of unifying approaches across methodological domains. Furthermore, while analyses of large, public corpora abound in existing literature, limited work has been done to explore the relationship between of *private* textual communications, personality traits, and symptoms of mental illness. We present a framework for constructing rich feature spaces from digital text communications. We then demonstrate the efficacy of our framework by applying it to a dataset of private Facebook messages in a college student population ($N = 103$). Our results reveal key individual differences in temporal and relational behaviors, as well as language usage in relation to validated measures of trait-level anxiety, loneliness, and personality. This work represents a critical step forward in linking features of private social media messages to validated measures of mental health, wellbeing, and personality.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**; • **Applied computing → Psychology**.

Additional Key Words and Phrases: machine learning; social media; language; mental health; text mining

Authors' addresses: Sanjana Mendu, University of Virginia, Charlottesville, VA USA, sm7gc@virginia.edu; Anna Baglione, University of Virginia, Charlottesville, VA USA, ab5bt@virginia.edu; Sonia Baee, University of Virginia, Charlottesville, VA USA, sb5ce@virginia.edu; Congyu Wu, University of Texas at Austin, Austin, TX USA, congyu.wu@austin.utexas.edu; Brandon Ng, University of Richmond, Richmond, VA USA, bwn7zb@virginia.edu; Adi Shaked, University of Virginia, Charlottesville, VA USA, as5fz@virginia.edu; Gerald Clore, University of Virginia, Charlottesville, VA USA, gc4q@virginia.edu; Mehdi Boukhechba, University of Virginia, Charlottesville, VA USA, mob3f@virginia.edu; Laura Barnes, University of Virginia, Charlottesville, VA USA, lb3dp@virginia.edu.

2573-0142/2020/10-ART144 $15.00
https://doi.org/10.1145/3415215

ACM Reference Format:
Sanjana Mendu, Anna Baglione, Sonia Baee, Congyu Wu, Brandon Ng, Adi Shaked, Gerald Clore, Mehdi Boukhechba, and Laura Barnes. 2020. A Framework for Understanding the Relationship between Social Media Discourse and Mental Health. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 144 (October 2020), 23 pages. https://doi.org/10.1145/3415215

1 INTRODUCTION

Digital text communications (DTCs) exchanged over social networks such as Facebook, Twitter, and Instagram and platforms such as Facebook Messenger, Twitter, and WhatsApp form the the collective touchstone of modern communication for young adults. Recent work suggests that the use of DTC platforms provides unique insight into the mental health and well-being of young adults [7, 84, 96] and may strongly influence mental health outcomes. Variations in perceptions of social support on Facebook, for instance, have been found to be associated with depression [66], and Facebook use in general has been associated with declines in well-being over time among college students [53]. Though the role of DTC platform use in mental health remains disputed, these tools have been found to facilitate access to informational and social support [29], especially for groups who may struggle to obtain support in the wild. Studies have shown that DTCs have been associated with improved mental health outcomes in vulnerable populations, including breast cancer patients [4], individuals with severe mental illness (e.g., schizophrenia and bipolar disorder) [62], homeless youth [73], and young adults with diabetes [38]. Among college students, greater perceived support on platforms like Facebook has been tied to less stress [94] and greater physical and emotional wellbeing [34, 61].

Previous work has focused on several important domains within DTC analysis, such as the topological structure of social networks and the linguistic content of messages. Computational approaches to classifying and predicting mental health issues from DTC datasets vary widely by dataset, platform, and mental health issue. Some rely on tools such as Linguistic Inquiry and Word Count (LIWC) to identify common language features of different issues [28, 32, 65]. Others examine how variations in temporal communication patterns and social network topology influence symptom prevalence [13, 33]. Researchers typically use a combination of methods to explore DTCs across domains. However, to our knowledge, no guiding frameworks exist to facilitate analysis across DTC domains with respect to mental health. A few existing frameworks in the mobile sensing literature have included DTCs as key components of research in mobile sensing for mental health [1, 59]. However, DTCs are merely singular components of these frameworks, rather than the main focus. Moreover, existing works in DTC analysis for mental health have relied primarily on datasets of public and semi-public content from Facebook and Twitter. Limited analyses have been conducted on private messages (e.g., Facebook messenger datasets).

In this work, we present a feature extraction framework for DTC analysis and apply this framework to guide exploratory analyses of private DTCs from a college student population. Our work provides two main contributions: 1) Establish a unifying hierarchy for DTC feature extraction methods, and 2) Identify individual differences in anxiety, loneliness, and personality within a college student population, as determined by these features. First, we provide a brief overview of the related literature. Then, we present our framework and explain how our feature extraction recommendations align with the related literature. Finally, we discuss an application of our framework to a private DTC dataset and highlight important findings afforded by our comprehensive approach.

2 RELATED WORK

As social media platforms have grown to form the foundation of modern digital communication, DTC datasets have proliferated. These exchanges comprise a rich corpus of interpersonal exchanges, which can provide insight into how mental health issues manifest in different social contexts. Further, because social media data is recorded in the present by an individual, it also serves as a complementary verbal sensor to understand the psychological dynamics of an individual, beyond non-verbal passive sensors [77]. In the following section we discuss the value of DTCs within feature extraction frameworks. We then highlight the diversity of existing approaches for extracting features from DTC data. Finally, we argue that unifying these methods within a novel framework can lead to comprehensive, multi-faceted insights into the manifestation of mental health in daily life.

DTCs comprise a rich corpus of interpersonal exchanges. Together, the features extracted from these exchanges provide insight into mental health status, wellbeing, and personality, and how these manifest in different social contexts. Previous work has contextualized DTCs as their own kind of sensor streams which complement passive sensing technologies [77]. Often, researchers have included DTCs as components in comprehensive frameworks for mobile sensing. For instance, both Mohr et al. and Abdullah and Choudhury's frameworks mapped raw sensor data (including DTCs such as SMS messages) to higher-level features and to mental health states [1, 59]. Further, Aung et al. [5] and Burns et al. [17] presented tripartite frameworks for sensed data which allow for the inclusion of "soft sensors," such as call and text logs. While these and other existing frameworks have used mobile sensing to understand mental health in context, none, to our knowledge, have focused exclusively on DTCs. By focusing exclusively on DTCs, we introduce an opportunity to extract richer social contexts and improve our understanding of their role in mental health, wellbeing, and personality.

DTC lexica have been shown to reflect individual communication styles and provide insight into personal traits, relationship quality, and mental state. Researchers have identified shared vocabularies and interpersonal differences in message semantics among individuals with mental health issues [12, 24, 25, 48, 65, 85]. For example, Coppersmith et. al showed that a character language model can discriminate among mental health issues, meaning that "spaces, punctuation, and emotico[n] usage" differs by condition [25]. Linguistic Inquiry and Word Count (LIWC) has proven popular among psychologists and HCI researchers alike for its ability to uncover links between personality, language, and mental health issues. LIWC analysis has been used to predict personality traits [37, 79, 81], emotion [68], and such conditions as depression [27, 32], suicidality [28, 65], and disordered eating [90]. Sentiment analysis is another popular method for characterizing on-line textual expression [25, 68, 89, 91]. Alternatives to closed-vocabulary method include unsupervised, open-language approaches, such as topic modeling (i.e., latent dirichlet allocation (LDA)) [32, 79, 81], and word embeddings [10, 89]). These techniques are used to extract textual patterns that describe the relationship between different linguistic structures and their effect on the overall meaning of a given text. By examining both the syntax of messages and the context within which an individual is communicating, researchers uncover data-driven language structures rather than rely on pre-defined vocabularies.

Researchers have explored communication patterns in different temporal contexts, including daily [3, 65] weekly [33], and multi-month contexts [12, 13]. Patterns have included communication around situational events [72], as well as communication frequency overall [15, 37, 68] and during different epochs [37]. Burke and Kraut, for example, used temporal and topological properties to understand social processes on Facebook following a job loss [13]. Researchers have also leveraged social network analysis methods to construct graphical structures of DTC data, abstracting individuals as nodes and their communications as edges [78, 95]. Relational patterns can be similarly inferred by constructing graphical networks from a dataset of directed messages. From these networks, researchers have found important links between structural patterns (e.g., network size [15, 37, 79, 90], betweenness [37], density [37], transitivity [37, 91], tie strength [15, 91], group associations [6, 15], persistence of social signature [21], turnover [21], rank dynamics [21, 68], interaction diversity [90]) and a diverse range of mental health issues.

Employing a combination of these methodologies could reveal insights about how, why, and when symptoms manifest in digital communications and could help researchers transcend traditional disciplinary boundaries. In this paper, we present a framework for extracting features from digital text communication datasets that draws from diverse methodological approaches across research domains and provides an avenue for logically deconstructing DTC datasets.

3 SOCIALTEXT: A FRAMEWORK FOR EXTRACTING FEATURES FROM DTCS

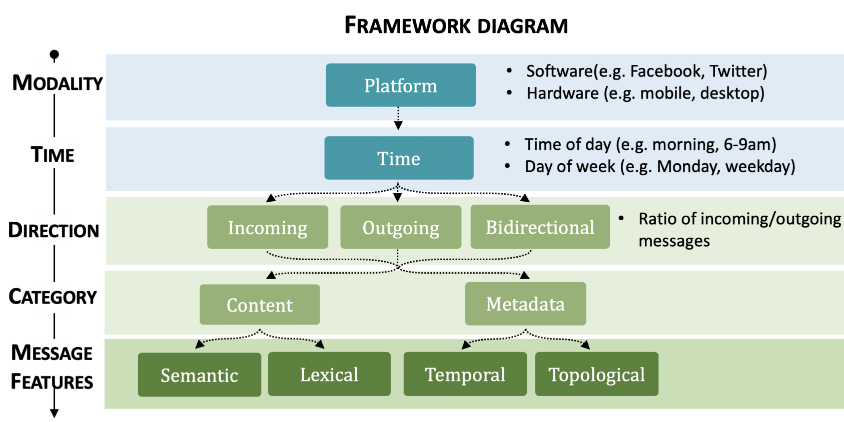

Fig. 1. Visual representation of SocialText framework

To effectively identify and analyze the relationship between underlying social contexts and mental health issues evidenced in digital text communication data, researchers must extract a comprehensive corpus of features from raw textual data streams. To this end, we propose *Social-Text*, a framework that unites diverse methodological approaches to analyzing the relationship between DTCs and mental health issues. The goal of the *SocialText* framework is to provide a clear, comprehensive method for creating informative, organized feature spaces for analysis of DTC social semantics. Figure 1 provides a visual overview of *SocialText*. In the following section, we discuss the relevance of each of the framework's layers to social context and mental health states.

Modality pertains to both the software and hardware used to send and receive communications. A unique modality is defined in terms of the software platform (e.g., Facebook, SMS) and/or device used (e.g., laptop, phone). Identifying an appropriate modality for the research question at hand is of critical importance, as messaging behaviors vary across different platforms. For instance, Facebook messaging behavior (e.g., commenting on friends' posts) has been shown to be a strong estimator of social ties [14] and has been studied extensively in work focused on social support [13, 15].

Time refers to the time window of interest (i.e., hour, day, week) for analysis. Previous analyses of sensed data streams rich in mental health data have highlighted the importance of determining the appropriate time window for feature extraction. For instance, Loveys et al. were able to identify distinct affective micropatterns in Twitter posts using three-hour-long time windows [57]. Further, Saeb et al. examined factors such as the relationship between GPS visit duration and depression [76], and Doryab et al. extracted features such as call logs during different epochs (6 hour time windows) [30]. Interestingly, few previous works have focused on the underlying social context tied to the timing of DTCs. We argue that the time at which individuals send and receive messages can reveal much about interpersonal relationships and communication styles.

Appropriate time windows vary depending on the outcome variable being studied. Consider, for instance, a study examining *state* anxiety vs. *trait* anxiety. State anxiety is influenced by the user's current context or environment and tends to be more dynamic, while trait anxiety is inherent to the individual and changes slowly if at all over time [36]. Barring any significant disruptions in routine, the number of messages an individual sends in a week is likely to remain relatively constant and thus follow an individual's baseline trait anxiety, while daily messaging patterns are likely to experience greater volatility and thus follow state anxiety. Moreover, an individual may shift from a pattern of consistent engagement with her social circle on weekends (when she has more free time) to short episodes of high engagement throughout the week (when she is busy with work or school) with prolonged lapses after each episode. While these patterns may appear similar in an aggregated week-level measure, analyses of daily messaging rates may reveal granular communication patterns in flux and may provide evidence of fluctuations in an individual's mental state.

Direction comprises three distinct categories of messages: solely *incoming*, solely *outgoing*, or *bidirectional* (i.e., conversations as a whole). Message direction has garnered much interest among researchers studying social connections online. For instance, Burke et al. highlighted how inbound directed communications (messages sent between individual friends in a social network) were predictive of bridging social capital [14]. Further, Burke et al. found that directed communication from strong ties (e.g., close friends) moderates stress in those who have lost a job [13]. Notably, related works exploring mental health expression via DTCs have tended to focus on linguistic [48] and time-based features [31] as opposed to message direction. We argue that a deeper dive into analysis of message direction can reveal egocentric aspects of the underlying social context of a conversation (e.g., who dominates the conversation). Further, such an exploration is worthwhile for identifying patterns linked to certain personality characteristics (e.g., extraversion) or mental health states (e.g., anxiety). Outgoing message features, in particular, can reveal relationships between an individual's communication practices and their mental state. For example, loneliness and depression have been associated with withdrawal and isolation [60, 64]. Using *SocialText*, researchers could examine whether users who are depressed or lonely at baseline send fewer outgoing messages, on average. Further, bidirectional message features that describe all messages irrespective of whether they are incoming or outgoing, reveal factors such as discussion quality and conversation dynamics (e.g., who is talking more).

Category distinguishes between two distinct categories of features: *content* and *metadata*. Content features comprise linguistic features such as verb or pronoun frequency. Content features are useful for identifying shared vocabularies and interpersonal differences in message semantics between members of a social network. Metadata features, meanwhile, comprise features such as the timing and frequency of message exchanges and the overarching network structure. Both content and metadata features have been used in tandem throughout the literature to characterize mental health conditions and to predict onset of a condition [26–28]. We define an additional layer of **Message Features** within each of these categories that addresses the different features that can be enumerated within them. This layer does not further partition the data, but rather enumerates the aggregated features that can be calculated based on individual messages. Content features comprise the following message feature layer components: *lexical* features, which refer to vocabulary and term-related qualities of message content, and *semantic* features, which capture the relationships between words within a set of messages and their effect on the overall tone and meaning of the communication. Metadata features, on the other hand, are broken down into *temporal* features, which refer to time-sensitive message characteristics, and *topological* features, which refer to social network structures.

4 STUDY DESIGN

In this section, we present an application of the *SocialText* framework to a dataset of private Facebook messages collected from a sample of college undergraduates at a U.S. university. We examine the relationship between social media usage and mental health at the individual/trait level. By understanding the social strategies that people use in their everyday life, and whether different strategies may be most effective for people with different psychological traits and mental health issues, we hope to achieve a better understanding of mental health for all.

4.1 Participants

Participants ($N = 103$) were recruited from undergraduate psychology classes at our university and received course credit as compensation. By recruiting young adults in a university setting, we obtained a relatively homogenous sample with respect to psychosocial stressors and life experiences, thereby eliminating many potential "nuisance factors". Our population was evenly sampled with respect to gender, with 51 female participants and 52 male participants. Participants' ages ranged from 18-22 years old, with the average age being 19 years old.

4.2 Psychological Measures

To assess participants' mental state, we administered clinically validated measures of anxiety, loneliness, and personality during an initial in-laboratory session. Each of the measures described below has been previously studied in a trait-level context [42, 83, 86].

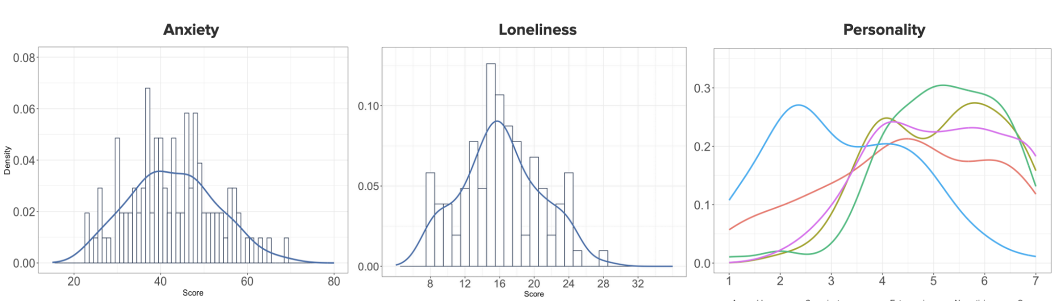

Fig. 2. Distribution of Anxiety (M = 42.77; SD = 9.95), Loneliness (M = 16.16; SD = 4.57), and Personality Trait [*Openness*: (M = 5.13; SD = 1.31), *Extraversion*: (M = 5.18; SD = 1.16), *Agreeableness*: (M = 4.43; SD = 1.69), *Neuroticism*: (M = 3.21; SD = 1.39), and *Conscientiousness*: (M = 5.14; SD = 1.22)] levels among the participants

Anxiety: The State Trait Anxiety Inventory (STAI; [86]) assesses two distinct dimensions of anxiety: (1) *state anxiety* (a temporary condition resulting from an individual's current state) and (2) *trait anxiety* (a long-standing quality of the individual). In this analysis, we consider trait anxiety to be our proxy for anxiety on the individual level. Participants rated the degree to which they generally identified with each statement (e.g., "I feel satisfied with myself") from 1 ("almost never") to 4 ("almost always").

Loneliness: The UCLA Loneliness Scale (ULS-20; [75]) is a widely used loneliness measure. We used an alternative short-form measure in this study (ULS-8; [47]). Participants rated the degree to which they generally identified with each statement (e.g., "I feel isolation from others") from 1 ("I never feel this way") to 4 ("I often feel this way").

Personality: The Ten-Item Personality Inventory (TIPI; [42]) provides measures of the "Big Five" (i.e., Five-Factor Model) dimensions of personality: *Openness*, *Extraversion*, *Agreeableness*, *Neuroticism*, and *Conscientiousness*. Participants rated the degree to which they agreed with each statement, specifically the extent to which each pair of traits applied to them (e.g., "extraverted, enthusiastic"), from 1 ("disagree strongly") to 7 ("agree strongly").

4.3 Facebook Messages

We requested that participants provide us with their Facebook messages since the time of account creation and, optionally, their public Facebook logs. Participants who opted not to provide us with their logs still received full credit for participating in the study. Those who opted to provide their logs downloaded them from the Facebook website during an in-laboratory session. Due to the lack of download configuration options available at the time the logs were downloaded, logs dated back to the creation of the account.

To account for individual differences in account creation date, we calculated the number of days of available data for the participant using the most recent account creation date ($T \approx 5$ months) and used that as a uniform time interval to compare all participants fairly. Overall, the dates used in this analysis span from June 8 to November 7, 2016. All data falling outside this specified time range were omitted from the current analysis. Our final dataset comprises 1,051,858 messages across all participants, with an average of 10,212 messages (σ = 27,869) and 48 unique chats (σ = 37) per participant.

4.3.1 Ethical Considerations. Aggregation of private data into large, readily-available datasets has come under intense scrutiny in the wake of events such as the Cambridge Analytica scandal. Though the debate over the extent to which private information may be ethically collected continues, ethical researchers agree that participant privacy must take utmost precedence in all studies involving sensitive data. We took careful steps to protect participants' privacy at each step of the research process. Participants signed a consent form at the beginning of the study and a material release form at the end of the study. A member of the research team was present for all lab sessions to explain the consent process and to answer the participants' questions.

This study required the use of private data for several reasons. First, the public and private selves are often quite different, especially with regard to DTCs. Free disclosure of mental health concerns in public online spaces (e.g., public Tweets and Facebook posts) may be met with lack of response from one's network due to the hypothesized "positivity bias" against negative status updates [97]. We hypothesized that private DTCs are more likely to contain naturalistic mental health information. Moreover, when users feel able to discuss health concerns privately, the quantity of messages and thus the size of the dataset should increase. Having more data allows for more accurate observations about DTC communication patterns, such as density of messages by time of day and how it relates to personality and mental health.

5 MODELING PSYCHOLOGICAL TRAITS FROM DTCS

5.1 Feature Extraction

In accordance with the *SocialText* framework structure, the features we extracted cover a broad range of DTC properties which we divide into four distinct categories: *Semantic*, *Lexical*, *Topological*, and *Temporal*. Table 1 provides a comprehensive list of extracted features [1].

Feature Domain	Feature Name	#	Direction
Lexical	LIWC	184	↑↓
Semantic	TF-IDF	6,348	↓
	LDA Topic Usage	100	↓↑
Temporal	Latency	2	↑↓
	Hourly Proportion	72	⇕↑↓
Topological	Number of Individual Alters	3	⇕↑↓
	Number of Group Alters	3	⇕↑↓
	Maximum Edge Weight	3	⇕↑↓
	Entropy of Edge Weights	3	⇕↑↓
	Mean/SD Persistence	6	⇕↑↓
	Mean/SD Turnover	6	⇕↑↓

Table 1. List of features. Direction (⇕: bidirectional, ↑: outgoing, ↓: incoming)

5.1.1 *Semantic.*

Semantic features describe the relationship between linguistic structure and the meaning of a given text. We created a set of linguistic structures in the corpus using the Natural Language Toolkit (nltk) TweetTokenizer [56] to split each message into unigrams. We also extracted bigrams and trigrams (e.g., phrases) - two and three-word sequences that occur at rates much higher than chance (e.g., "happy birthday", "I love you") - by calculating the pointwise mutual information (PMI) [23, 54] of each phrase (i.e., a ratio of the joint-probability to the independent probability of observing the phrase within the aggregated corpus of messages):

$$\text{PMI(phrase)} = log \frac{p(\text{phrase})}{\prod_{w \in \text{phrase}} p(w)} \tag{1}$$

We retained all bigrams and trigrams with PMI values greater than 3 times the number of words in the phrase. The resulting vocabulary consisted of 6,348 words and phrases. To reduce the number of features, we kept words and phrases that were used at least once by at least 10% (n=10) of the population. We calculated the *Term Frequency - Inverse Document Frequency (TF-IDF)* of each term in the vocabulary described above in order to measure each term's usage within each participants' set of messages. TF-IDF serves as a useful measure for between-subjects analyses such as ours because it accounts for the relevance of terms across multiple documents.

We also identified *topics* - clusters of frequently co-occurring words in our corpus - using Latent Dirichlet Allocation (LDA) [11]. The generative LDA model assumes that documents (i.e., a participant's complete set of private Facebook messages) contain a combination of topics, and that topics are a distribution of words (i.e., observations) for which the latent variables can be estimated through Gibbs sampling [44]. For this analysis, we leveraged the implementation of this algorithm provided in the Mallet package [58] to produce 100 naturally-occurring topics, each consisting of many words with relative weights. We then calculated each individual's use of each topic, defined

[1] Our feature extraction code is shared as an open source tool at https://github.com/BarnesLab/SocialText-Feature-Extraction

as the probability of using a topic:

$$p(\text{topic}|\text{user}) = \sum_{\text{word}\in\text{topic}} p(\text{topic}|\text{word}) * p(\text{word}|\text{user}) \tag{2}$$

where $p(word|user)$ is the individual's normalized word use.

5.1.2 Lexical. DTC lexica reflect individual communication styles and provide insight into personal traits, relationship quality, and mental state, among other factors. We extracted lexical features using the popular *Linguistic Inquiry and Word Count (LIWC)* method, which has been rigorously validated in the context of psychometric analysis of textual data [88].

5.1.3 Topological. The topology of the ego-centric network formed by an individual's social circle can provide significant insight into the individual's personality traits [87]. Suppose, over a time period (5 months in this study), a subject (ego) exchanged (i.e.„ sent and/or received) at least one message with K unique alters. The K alters can be partitioned into K_1 individual alters representing individual recipients and K_2 group alters representing two or more recipients giving $K = K_1 + K_2$. We extracted the following features to capture the size of individuals' social networks: *number of individual alters* K_1 (i.e.„ the number of contacts representing an individual with whom a subject exchanged at least one message); and *number of group alters* K_2, (i.e.„ the number of contacts representing at least two people with whom a subject exchanged at least one message).

The messages exchanged between the subject and an alter constitute the edges in the network, and we define *edge weight* as the proportion of messages exchanged with an alter (individual or group) among all alters. We denote as p_r $r \in \{1, \ldots, K\}$ the r-th highest proportion of messages exchanged with an alter among all alters, and the distribution of proportions/edge weight over all alters as $P = \{p_1, \ldots p_K\}$. We extracted the following features to capture differences in exchanges in the context of an individual's social network: *entropy of edge weight* $H(P) = -\sum_{p\in P} p \log(p)$ (i.e.„ the Shannon entropy of the proportions of messages exchanged with all alters a subject had). This measure quantifies how a subject distributes their time across multiple threads of conversations; and *maximum edge weight* p_1 (i.e.„ the proportion of messages exchanged with the alter with whom the subject exchanged the most messages).

We also sought to characterize the variation of social dynamics over more granular time intervals. We calculate two measures: the persistence of social signatures and the turnover in ego-centric networks. These measures come from existing work on ego-centric network dynamics [2, 21], which proposed and applied these measures to phone call and Bluetooth encounter networks. To calculate, we first divide the 5-month observation period into 21 week-long periods $\{w_1, \ldots, w_{21}\}$. For each pair of consecutive periods (w_i, w_{i+1}) $\forall i \in \{1, \ldots, 20\}$ we calculate the following features: (1) *persistence of social signature*, defined as the Jensen-Shannon divergence between the P's calculated from w_i and w_{i+1},

$$\text{persistence}(w_i, w_{i+1}) = H\left(\frac{P_{w_i} + P_{w_{i+1}}}{2}\right) - \frac{H(P_{w_i}) + H(P_{w_{i+1}})}{2}, i \in \{1, ..., 20\} \tag{3}$$

; and (2) *turnover of ego-centric network*, defined as the Jaccard difference between the two sets of alters, $A(w_i)$ and $A(w_{i+1})$, corresponding to w_i and w_{i+1} for a subject, concretely:

$$\text{turnover}(w_i, w_{i+1}) = \frac{|A(w_i) \cap A(w_{i+1})|}{|A(w_i) \cup A(w_{i+1})|}, i \in \{1, ..., 20\} \tag{4}$$

We obtain 20 values for each measure and calculate the mean and standard deviation, producing 4 features in total: mean persistence, standard deviation of persistence, mean turnover, and standard deviation of turnover.

5.1.4 Temporal. The time at which individuals send and receive DTCs can reveal much about underlying social context, including interpersonal relationships and communication styles. We calculated the *hourly distribution* of messaging activity (i.e., the proportion of messages sent during each hour of the day) from the aggregated collection of each participants' Facebook message logs. We also calculated *latency* for both outgoing and incoming messages, where outgoing latency is the average amount of time (in minutes) that a participant takes to respond to a message they receive, and incoming latency is the average amount of time (in minutes) it takes for a participant to receive a response to a message they sent.

5.2 Predictive Modeling

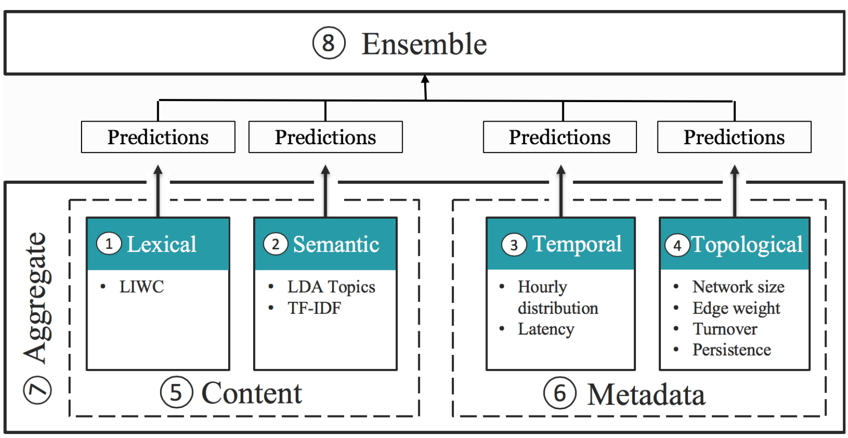

Fig. 3. The above figure provides a visual representation of our modeling process

We began our evaluation by testing predictive models for each feature category independently. To reduce the effect of irrelevant features and mitigate the curse of dimensionality, we used a random forest classifier to select a subset of the 10 most relevant features to the given outcome for each feature domain independently, based on the mean decrease in Gini impurity when a feature is used to partition the data. We then use a Support Vector Machine (SVM) and leave-one-subject-out cross validation (LOSOCV) to predict a binary classification for each psychological measure using unique groupings of the messages features as input.

We also investigated the effect of combining the content feature spaces (i.e., Semantic & Lexical) and metadata feature spaces (i.e., Temporal & Topological) on model performance. Finally, we used two approaches to combining features across all four message feature domains: ensemble and aggregated. For the ensemble model, we used stacked generalization [93] to predict psychological characteristics. This approach is advantageous because it overcomes the potential for features from larger domain spaces (i.e., Semantic & Lexical) to overpower smaller domain feature spaces (i.e., Temporal & Topological), since the representation of knowledge from each domain is condensed in the form of each independent model's prediction. For the aggregated model, we combined features across message feature domains into a single feature space. We then applied the same Random Forest approach used for the independent domain models to reduce the dimensionality of the cumulative feature space.

6 RESULTS

6.1 Temporal

Surprisingly, outgoing latency was *not* a discriminating feature with respect to anxiety, loneliness, or personality. On the other hand, incoming latency was one of the more important temporal features for predicting four out of the seven psychological measures. As shown in Figure 4, anxious and lonely individuals' friends took longer to respond to them (i.e., anxious and lonely populations' communications exhibited a greater incoming latency). Furthermore, extroverts and introverts took about the same amount of time to respond to messages they received, on average. However, introverts' friends took longer to respond to them than did extroverts' friends.

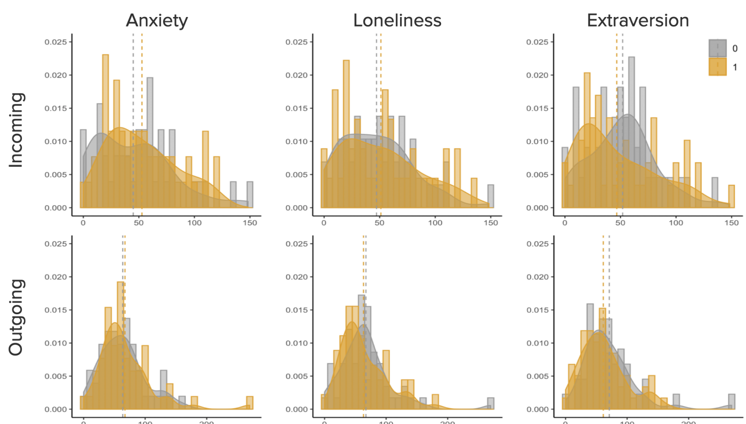

Fig. 4. Difference in incoming and outgoing latency between "high" (i.e., 1) and "low" (i.e., 0) anxiety, loneliness, and extraversion groups

Individual differences in psychological attributes also moderated when participants engaged in conversations on Facebook Messenger. Anxious participants showed notable variation in evening DTC activity compared to non-anxious individuals, especially between the hours of 9pm and 12am. More specifically, anxious participants sent more messages at 9pm and 11pm and received more messages at 10pm than non-anxious participants, as highlighted in Figure 5. Lonely participants

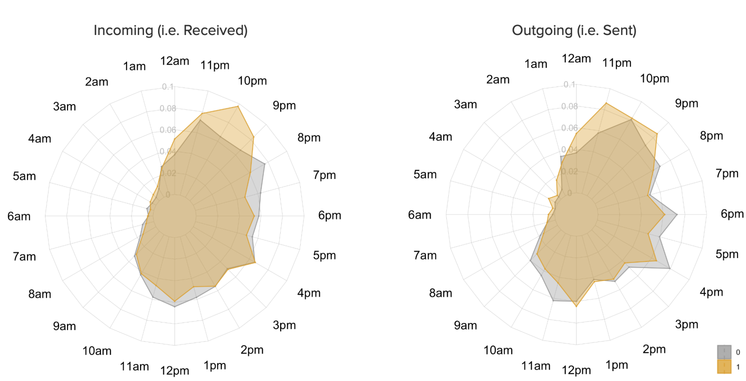

Fig. 5. Comparison between "high" (i.e., 1) and "low" (i.e., 0) anxiety classes with respect to the average proportion of messages received/sent during each hour of the day

exhibited a similar divergence, although less consistently and during a slightly shorter time window (9pm to 11pm). These results suggest a marked diurnal shift in communication patterns in our anxious participant population. Moreover, participants exhibited notably different patterns of Messenger usage at the beginning and end of the standard work day (i.e., 8am, 5pm) when compared to emotionally stable participants.

6.2 Topological

Turnover in an ego-centric network and persistence of social signature were found to be important factors within the aggregate models for neuroticism. Specifically, individuals with high vs. low levels of loneliness exhibited noticeably different levels of persistence in general. This suggests that trait loneliness moderates the extent to which participants engage in consistent messaging behavior over the course of the five-month interval we studied. Neurotic and emotionally stable individuals exhibited a similar divergence in interpersonal dynamics as measured by average bidirectional persistence social signature. This divergence suggests that neuroticism moderates the consistency and variability of participants' digital social circle over the five-month interval.

Maximum edge density measure proved to be an effective proxy for biased communications (i.e., concentrating messages primarily within a single chat) within our population. Participants in the "high" and "low" agreeableness groups were characterized by differences in maximum outgoing edge density, while participants in the "high" and "low" conscientiousness groups were characterized by differences in maximum incoming edge density. Notably, participants in the "high" and "low" loneliness group were characterized by differences in both incoming and outgoing edge weight entropy. This indicates that individuals' level of loneliness mediates the consistency of their messaging behavior.

The number of alters also proved to be a valuable discriminator within the metadata models for extraversion and conscientiousness, contributing to an improvement in model performance 0.042 and 0.198 respectively when compared to the independent models. The overall number of alters over the five-month interval was found to be one of the more discriminating features of extroverts compared to introverts. This trend extends to individuals with "high" vs. "low" levels of loneliness. Participants in the "high" and "low" conscientiousness groups differed in the number of individual alters to which they sent messages with conscienctious individuals sending messages to more alters than less conscientious were sent.

6.3 Lexical

Anxious individuals used more first and third person plural pronouns, relativity (spatial, temporal), and male references than less anxious individuals, as evidenced in Figure 6. Participants' levels of anxiety were also predictable by the authenticity of language used by those in their network, as well as incoming content containing relativity, netspeak, certainty, and informal language.

Levels of loneliness were distinguishable by participants' outgoing language. Lonely individuals used more words in messages they sent to others and , used more 3rd person plural pronouns, periods, and adverbs. They also differed in their discussion of friends and other affiliations. Discussion related to certainty and interrogative topics, as well as level of authenticity in outgoing language further distinguished the two groups.

Extroverts and introverts (i.e., participants in the "high" and "low" extraversion groups) varied in their use of personal pronouns and discussion of social processes, as well as their social contacts' use of function words, exclamation marks, netspeak, personal pronouns, pronouns, achievement, 1st person singular pronouns. Participants in the "high" and "low" conscientiousness groups were more readily discriminated by the content of messages individuals received vs. sent. Specifically,

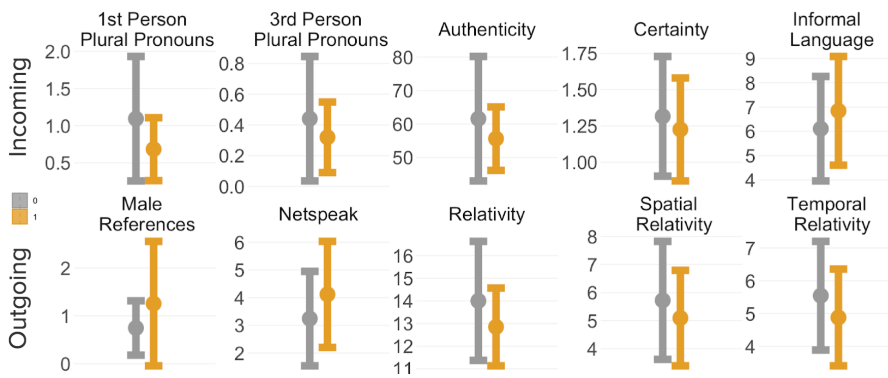

Fig. 6. Difference in incoming and outgoing message content between "high" (i.e., 1) and "low" (i.e., 0) anxiety classes in terms of LIWC measures

participants in the "high" and "low" conscientiousness groups were discriminated by use of informal language, punctuation, netspeak, and words longer than 6 letters contained in messages they received within the independent lexical model. Incoming message language pertaining to assent, affiliations, past focus, risk, and drives was also used to discriminate between participants in the "high" and "low" conscientiousness groups. This pattern suggests that participants' levels of conscientiousness mediate the formality of language used by their social contacts on Facebook Messenger. Surprisingly, the extent to which participants discussed biological processes was an important discriminating factor between "high" and "low" groups on both extraversion and agreeableness.

6.4 Semantic

As mentioned in the *Model Performance* section, the independent semantic models outperformed the other independent models on five out of the seven psychological measures. This outcome suggests that the semantic features we extracted from the private Facebook message corpus were the best predictors of individual differences in personality traits and mental health issues within this college student population. Differences in topic usage across the "high" and "low" groups for the psychological measures yielded a number of interesting findings. Table 2 provides examples of meaningful topics we extracted from our corpus.

Discussion of Pokémon GO (i.e., topic ID = 46), a mobile game released in the United States on July 6, 2016 [92], emerged as a meaningful discriminator for anxiety, loneliness, and neuroticism. Discussion of sports (i.e., topic ID = 61) emerged as a meaningful discriminator for anxiety and neuroticism, supporting the idea that engaging in physical activity plays a key role in college students' emotional stability. Discussion of spiritual music (i.e., topic ID = 78) also emerged as a meaningful discriminator for anxiety. Use of words related to social support (i.e., topic ID = 88) provided meaningful context for differentiating participants with "high" and "low" levels of agreeableness. Discussion of alcohol and partying (i.e., topic ID = 90) was a meaningful for differentiating individual with "high" and "low" levels of openness. Extroverts and introverts (i.e., individuals with "high" and "low" levels of extraversion) exhibited notable differences in their discussion of emotional processes (i.e., topic ID = 44) via private messages on Facebook. Furthermore, extroverts and introverts also varied in their use of words related to recreation (i.e., topic ID = 41), which may reflect existing psychological associations between extraversion and positive affect.

Topic Label	ID	Terms
Election	40	trump, vote, election, president, voted, people, america, voting, rip, pretty, debate, votes, wins, country, politics, everyone, years, republican, winning, scared
Recreation	41	omg, haha, night, weekend, going, week, last, next, back, fun, come, coming, hahaha, dude, man, yay, party, visit, meet, nice
Emotional Processes	44	feel, talk, things, like, really, sorry, know, okay, think, someone, person, time, talking, life, something, sad, tell, anything, felt, miss
Pokémon GO	46	pokemon, caught, level, team, anyone, catch, gym, tonight, game, walk, find, house, found, mystic, around, valor, blue, playing, egg, people
Sports	61	game, play, team, playing, last, yeah, played, ball, soccer, games, beat, though, hit, football, hard, basketball, lost, time, damn, pretty
Spiritual Music	78	song, listen, music, man, time, pretty, could, never, high, every, family, without, made, makes, ever, times, though, great, different, god
Social Support	88	thanks, thank, great, miss, school, day, work, best, luck, fun, well, send, class, excited, aww, awesome, start, uva, summer, already
Alcohol	90	drink, party, drunk, drinking, night, alcohol, drinks, parties, people, beer, tonight, fun, sober, getting, drank, frat, shots, boy, wine, gone

Table 2. Topics present in private Facebook messages

		Anx.	Lon.	Extra.	Agree.	Open.	Neuro.	Consc.
Message Features	Lexical	**0.765**	0.678	0.627	0.547	**0.762**	**0.667**	0.615
	Semantic	0.701	**0.732**	**0.717**	**0.634**	0.734	0.594	**0.716**
	Temporal	0.698	0.705	NA*	0.580	0.548	0.641	0.500
	Topological	0.239	0.684	0.413	0.522	0.486	0.636	0.376
Category	Content	**0.694**	**0.743**	**0.660**	0.574	**0.796**	**0.646**	**0.687**
	Metadata	**0.694**	0.698	0.455	**0.611**	0.619	0.614	0.574
Combined	Aggregate	0.707	0.692	0.667	0.587	**0.789**	0.660	0.720
	Ensemble	**0.752**	**0.793**	**0.698**	**0.743**	0.774	**0.708**	**0.768**

Table 3. The above table shows each model's performance as measured by F1 score. * denotes an undefined F1 score resulting from a zero-valued recall and precision measure for the given model.

6.5 Model Performance

Independent Approaches: The semantic model outperformed all other independent models for predicting 4 out of the 7 psychological measures, with performance ranging from 0.634 to 0.732 overall. The lexical model performed better than the semantic model for predicting anxiety (F1 = 0.765), openness (F1 = 0.818), and neuroticism (F1 = 0.667) with a relative improvement of 0.064, 0.028, and 0.073 for each measure respectively. While the temporal model achieved moderate performance overall (Mean F1 = 0.524), it performs particularly poorly for predicting extraversion. Notably, extraversion is also the only measure for which the topological model (F1 = 0.413) outperforms the temporal model (F1 = NA). This is particularly interesting given the topological model showed poor performance overall (Mean F1 = 0.479). This discrepancy suggests that interpersonal dynamics outweigh temporal factors with respect to characterizing extraversion as manifested in private social media discourse.

Content vs. Metadata: The content-based model, which used only lexical and semantic features as predictors, outperformed the metadata-based model for predicting the majority of the psychological measures. This result is supported by findings from prior studies that used only content features to predict personality traits [79, 81]. Notably, the metadata model performed better than the content model for predicting agreeableness. Additionally, the content model performed relatively poorly for predicting agreeableness (0.574) as compared to the other psychological measures (0.646 to 0.796), suggesting that agreeableness relates less to what people say and more to when and with whom people engage in private social media discourse.

Combined Approaches: Additionally, our results highlight that using all feature domains to predict trait measurements outperforms the independent models. While the relative performance improvements vary from measure to measure, the average performance of the aggregate and ensemble models (0.689 and 0.749 respectively) exceeds the average performance of any given independent model (0.479 to 0.690). Furthermore, the ensemble model outperformed the aggregate model for predicting the majority of the psychological measures (6 out of 7). These results suggest that considering each of the different dimensions of DTC data (i.e., message features) in equal measure, rather than heavily weighting any given one, not only supports a more comprehensive consideration of underlying factors but also improves the performance of predictive modeling approaches.

7 DISCUSSION

7.1 Understanding Mental Health from DTCs

Semantic results reveal new insights about anxiety and neuroticism in relation to factors such as political unrest, social activities, social support, and even musical preference. The emergence of Topic 46 ("Election") is unsurprising, given that we collected baseline measures in early November 2016, but it nevertheless affirms the relevance of political tensions to college students' mental wellbeing. Hoyt et al. found evidence of increased negative affect and cortisol levels in a US young adult population around the time of the 2016 US presidential election [49], pointing to the significant detrimental effects the election had on young Americans' mental health. Whether this effect is unique to the 2016 election remains undetermined. Moreover, in the context of our work, we foresee an opportunity to investigate the relationship between political discussion, communication patterns, and short-term mental health outcomes in our population.

Discussion of games, both virtual (e.g., Topic 46: "Pokémon GO") and physical (e.g., Topic 61: "Sports") as predictors of such conditions as anxiety and neuroticism reveals much about the role of social games in mental health and wellbeing. Specifically, more anxious or neurotic participants talked about games less than less anxious or neurotic participants. Existing research supports our result that Pokémon GO served an important role in wellbeing and emotion regulation, particularly among college students For example, Kari et al. found that many participants reported using the game as self-treatment for helping with anxiety [52]. Our findings are novel, however, for neurotic individuals, who have been shown to use variants of the words "depressed" and "lonely" more often [80], use anger words frequently in their posts, and use "social interaction words" more sparingly [37]. Neurotic individuals' lack of discussion around sports, which are naturally social activities, further solidifies this evidence that neurotic individuals may be socially isolated or withdrawn.

The relationship between "Social Support" (Topic 88) and agreeableness is also informative, given that agreeableness may be influenced by mental health symptoms. Social support, both perceived and tangible, has been shown to strongly influence mental health outcomes, both positively and negatively. For example, Grieve et al. found that connectedness on Facebook correlated with reduced anxiety and depression [43]. Further, Indian and Grieve found that greater perceived social support on Facebook was associated with higher subjective wellbeing among high-socially anxious users [50]. Additionally, Burke and Kraut showed that inbound directed communication (e.g., Facebook

Messenger texts, comments, wall posts) with strong ties is associated with increases in wellbeing, for Facebook users at large [16]. Future research with our dataset could investigate manifestations of social support beyond the Semantic domain, as well as the interplay between social support and personality trait expression. We also note that Topic 78 ("Spiritual Music") provides an opportunity for further exploration, given the great relevance of both spirituality and music (separately) to mental health outcomes [55, 82]

Knowledge of language patterns linked to mental health is often critical for early intervention and prevention of worsening symptoms. For example, De Choudhury et al. found that individuals who first posted to mental health subreddits, then later posted to a suicide-specific subreddit (r/SuicideWatch), had posts with poorer linguistic structure in general [28]. They also received fewer comments, a phenomenon De Choudhury et al. observed among suicidal individuals on Reddit [28]. Additionally, Eichstaedt et al. found that words associated with loneliness, hostility, and rumination were the strongest predictors of depression in medical records [32]. Further, previous LIWC analyses have shown that greater use of first person singular pronouns has been associated with depression (including postpartum depression) and suicidality [27, 28, 32, 67]. Interestingly, our lexical results showed that first person plural pronouns (e.g., "we", "us", "our") were discriminative for anxious individuals, but first person singular pronouns were not. This finding suggests the need for deeper exploration of the relationship between pronoun usage and mental health beyond depression and suicidality.

Our temporal results reveal that communication at late night hours is highly predictive of both loneliness and anxiety, with more lonely or anxious individuals communicating more during these hours than less lonely or anxious individuals. These have been shown to be comorbid with sleep disorders such as insomnia or hypersomnia [63], especially in young adults [18]. This pattern dovetails with existing literature documenting a general relationship between loneliness and adverse sleep outcomes [19, 20, 46]. The relationship between loneliness and adverse sleep outcomes has also been documented in adolescents [45]. Similar relationships have been found between those with anxiety disorders and diminished sleep quality [22]. In short, the present study lends support to the potential links among loneliness, anxiety, and sleep, which may manifest temporally through diurnal shifts in communication (e.g., lonely individuals being more likely to communicate at night). We also note that the anxious and introverted groups had longer inbound message latency, on average, which could indicate diminished engagement or involvement from their friend networks. Anxious individuals were especially prone to this phenomenon, as their friends used more filler words and netspeak. Diminished engagement has historically been found in depressed groups. For example, De Choudhury et al. observed lowered overall engagement among depressed individuals including mothers with postpartum depression (PPD); both groups tend to share less and interact less with their peers on Facebook [26, 27]. Moreover, the PPD group also exhibited sharp, sudden changes in their level of activity over time [26]. Collectively, our data and previous literature point to the need for deeper investigation into temporal communication pattern disruption and the relationship between latency and network engagement in depressed, anxious, and introverted populations.

Broadly, the exchange of DTCs within a social network represents an ever-shifting exchange of *social support*. Previous works have emphasized the importance of *directed communications*, a subset of DTCs, for accumulating social support (*social capital*) [14, 35] and increasing tie strength [15]. Social anxiety and loneliness have both been associated with having fewer friends on Facebook [39, 51], indicating a possibly hampered ability of socially anxious or lonely individuals to grow their online social support network. Moreover, both groups exhibit distinct styles of sharing personal information that may affect their ability to gather and retain social support.

The differences between high and low-loneliness individuals in topological characteristics, is also grounded in prior literature within social psychology pertaining to social exclusion. For instance, prior research suggests that the need to belong is so imperative to human existence [8] that humans have developed a complex regulatory system to re-establish belonging when it is threatened (deemed the Social Monitoring System [69]). For instance, prior research has shown that social exclusion motivates heightened attention and vigilance to others' social cues [70] in the service of repairing belonging. Interestingly, this relationship has also been shown for individual differences in levels of loneliness, such that higher amounts of loneliness are related to higher amounts of attention to social cues [40]. However, recent work has shown that rejection based on a stigmatized identity (hypothesized to be a more chronic form of rejection) causes decreased attention to social cues [74]. In this way, participants high in loneliness may have impaired motivation to re-affiliate, and thus demonstrate decreased persistence relative to participants low in loneliness. For example, Jin et al. showed that lonely individuals tend toward negative self-disclosure and less "communicating activity" (such as making comments on others' posts) [51], both of which could deter potential network friends from offering social support outright. Similarly, Fernandez et al found that individuals with greater social anxiety tend to include more information in their personal profiles [39], which may signal a need for validation through oversharing.

Combined feature domain models (i.e., aggregate and ensemble models) were nearly always more predictive of the validated psychological measures than were features from singular domains. These results demonstrate both the utility and importance of considering features from multiple domains when investigating links between DTCs, personality, and mental health.

7.2 Limitations

Our findings carry several limitations that should be addressed in future research. First, our study population was relatively small and homogenous ($N = 103$). Communication behaviors may be impacted by age and/or cohort effects, thereby limiting generalizability of our findings to an older population. Future work should look at a more diverse sample and examine whether age is a covariate. Second, our analysis focuses on private messaging patterns and does not account for public social media behavior and other social interactions (i.e., in-person, phone, SMS). By not accounting for social interactions on these other platforms, our experimental results and conclusions are biased toward private messaging behaviors on a specific type of platform (Facebook messenger). Furthermore, prior work suggests that texting behavior (e.g., sharing intense and private emotions) varies across different platforms [9]. Thus, our findings from private messaging patterns may not translate to dynamics on public-facing DTC platforms. Third, as evidenced by the poor performance of the topological predictive models (Mean AUC = 0.426), our measures of interpersonal dynamics were not as reflective of individual differences in psychological measures as we hypothesized. This is likely due to the egocentric, generalized manner in which we extracted these features from our dataset. A more detailed picture of social network dynamics beyond the egocentric properties (e.g., interactions between participants' social contacts) would allow for more effective characterization of the quality of interpersonal exchanges on this platform.

7.3 Implications for Future Work

Our analysis shows that the unifying structure of the *SocialText* framework intentionally highlights features that can be derived from DTC data and used holistically to identify social context in a way that facilitates better prediction of mental health outcomes from DTCs. While the upper layers define important variables for data partitioning, the lowest layer identifies categories of features that can be extracted from the messages themselves. Features pertaining to the semantics and lexicon of message content can characterize conversational context, while temporal and topological features can reveal social network ties and temporal messaging patterns. Considering all message features in combination provides a comprehensive characterization of the relationship between social dynamics of DTCs and participants' mental states, thus improving the performance of the resulting predictive models. Researchers can use *SocialText* to identify and leverage multiple methodologies for characterizing or predicting mental health states.

Ref.	Modality	Time	Category	Direction	Message Features	Health Outcome
[13]	Facebook	Months	Metadata	↓	Temporal, Topological	Stress, Social Support
[41]	Twitter	Month	Content	↑	Semantic, Lexical	Stress
[12]	Twitter	Months	Content	↑	Semantic, Lexical	Mood
[24]	Twitter	Multi-year	Content	↑	Semantic,Lexical	Depression, BPAD, PTSD, SAD
[65]	SMS	Day	Content	↓↑	Semantic, Lexical	Depression, Suicide
[3]	SMS	Day	Metadata	↑	Temporal	Communication Satisfaction
[33]	SMS	Week	Metadata	↓↑	Temporal	Depression
[71]	SMS	Month	Metadata	↑	Temporal	Social Anxiety, Loneliness
[48]	SMS	All times	Content	↑	Semantic	Neuroticism

Table 4. Existing Literature Table. Direction (↑: outgoing, ↓: incoming)

Table 4 provides a list of selected, relevant studies that utilize DTC data to study mental health outcomes. In this table, we map each study onto the *SocialText* hierarchy, demonstrating its flexibility. Moreover, this mapping highlights important methodological overlaps in the existing literature. For example, Elhai et. al. [33] studied depression with respect to temporal patterns in SMS data, while Nobles et. al. [65] studied the semantic and lexical features of a similar dataset. While these studies choose different time windows (or, rather, *Time* layer selections), they are similar along all other dimensions of *SocialText*'s structure. By using *SocialText* to identify similar studies, such as [33] and [65], researchers can streamline the process of creating new methodological approaches from the best aspects of existing approaches. Thus, *SocialText* facilitates the development of sound methodologies for mobile mental health sensing.

8 CONCLUSION

Analysis of digital text communications (DTCs) remains an open research area at the intersection of mental health and computing. DTCs are feature-rich characterizations of social context, yet remain largely underexplored in existing mobile sensing frameworks. In this paper, we have proposed the *SocialText* framework, which defines a hierarchical structure for holistically extracting features from DTC datasets. Features pertaining to the semantics and lexicon of message content can characterize conversational context, while temporal and topological features can reveal social network ties and temporal messaging patterns. Considering all message features in combination provides a comprehensive characterization of the effect of social dynamics of DTCs on participants' mental states and allows researchers to leverage DTC feature extraction methodologies across academic disciplines. Our results corroborate previously established results and reveal novel individual differences in temporal and relational behaviors, as well as in vocabulary usage and topics of discussion, on Facebook Messenger. This work provides a novel path forward for future analysis and discussion of the role of DTCs in personality, mental health, and wellbeing online.

9 ACKNOWLEDGMENTS

Research reported in this publication was supported by a University of Virginia Presidential Fellowship in Data Science and Hobby Postdoctoral and Predoctoral Fellowship Grant.

REFERENCES

[1] Saeed Abdullah and Tanzeem Choudhury. 2018-01. Sensing Technologies for Monitoring Serious Mental Illnesses. *IEEE MultiMedia* 25, 1 (2018-01), 61–75.

[2] Aamena Alshamsi, Fabio Pianesi, Bruno Lepri, Alex Pentland, and Iyad Rahwan. 2016. Network diversity and affect dynamics: The role of personality traits. *PLOS ONE* 11, 4 (2016), e0152358.

[3] Alexa Angster, Michael Frank, and David Lester. 2010-10. An Exploratory Study of Students' Use of Cell Phones, Texting, and Social Networking Sites. 107, 2 (2010-10), 402–404.

[4] Deanna J Attai, Michael S Cowher, Mohammed Al-Hamadani, Jody M Schoger, Alicia C Staley, and Jeffrey Landercasper. 2015. Twitter social media is an effective tool for breast cancer patient education and support: patient-reported outcomes by survey. *Journal of Medical Internet Research* 17, 7 (2015), e188.

[5] Min Hane Aung, Mark Matthews, and Tanzeem Choudhury. 2017. Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies: Aung et al. *Depression and Anxiety* 34, 7 (July 2017), 603–609.

[6] Yoram Bachrach, Thore Graepel, and Pushmeet Kohli. 2014. Your Digital Image: Factors Behind Demographic And Psychometric Predictions From Social Network Profiles (Demonstration). (2014), 2.

[7] Shrey Bagroy, Ponnurangam Kumaraguru, and Munmun De Choudhury. 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 - CHI '17*. ACM Press, Denver, Colorado, USA, 1634–1646.

[8] Roy F Baumeister and Mark R Leary. 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin* 117, 3 (1995), 497.

[9] Natalya N Bazarova, Yoon Hyung Choi, Victoria Schwanda Sosik, Dan Cosley, and Janis Whitlock. 2015. Social sharing of emotions on Facebook: Channel differences, satisfaction, and replies. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 154–164.

[10] Adrian Benton, Margaret Mitchell, and Dirk Hovy. 2017. Multi-Task Learning for Mental Health using Social Media Text. *arXiv:1712.03538 [cs]* (Dec. 2017). http://arxiv.org/abs/1712.03538 arXiv: 1712.03538.

[11] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3, Jan (2003), 993–1022.

[12] Johan Bollen, Huina Mao, and Alberto Pepe. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Fifth International AAAI Conference on Weblogs and Social Media*.

[13] Moira Burke and Robert Kraut. 2013. Using Facebook after losing a job: Differential benefits of strong and weak ties. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. ACM, 1419–1430.

[14] Moira Burke, Robert Kraut, and Cameron Marlow. 2011. Social capital on facebook: differentiating uses and users. In *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems - CHI '11*. ACM Press, Vancouver, BC, Canada, 571.

[15] Moira Burke and Robert E. Kraut. 2014. Growing closer on facebook: changes in tie strength through social network site use. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, Toronto, Ontario, Canada, 4187–4196.

[16] Moira Burke and Robert E. Kraut. 2016. The Relationship Between Facebook Use and Well-Being Depends on Communication Type and Tie Strength: Facebook AND Well-Being. *Journal of Computer-Mediated Communication* 21, 4 (July 2016), 265–281.

[17] Michelle Nicole Burns, Mark Begale, Jennifer Duffecy, Darren Gergle, Chris J Karr, Emily Giangrande, and David C Mohr. 2011. Harnessing Context Sensing to Develop a Mobile Intervention for Depression. *Journal of Medical Internet Research* 13, 3 (Aug. 2011), e55.

[18] Daniel J. Buysse, Jules Angst, Alex Gamma, Vladeta Ajdacic, Dominique Eich, and Wulf Róssler. 2008. Prevalence, Course, and Comorbidity of Insomnia and Depression in Young Adults. *Sleep* 31, 4 (2008), 473–480.

[19] John T Cacioppo, Louise C Hawkley, Gary G Berntson, John M Ernst, Amber C Gibbs, Robert Stickgold, and J Allan Hobson. 2002a. Do lonely days invade the nights? Potential social modulation of sleep efficiency. *Psychological Science* 13, 4 (2002), 384–387.

[20] John T Cacioppo, Louise C Hawkley, L Elizabeth Crawford, John M Ernst, Mary H Burleson, Ray B Kowalewski, William B Malarkey, Eve Van Cauter, and Gary G Berntson. 2002b. Loneliness and health: Potential mechanisms. *Psychosomatic medicine* 64, 3 (2002), 407–417.

[21] Simone Centellegher, Eduardo López, Jari Saramäki, and Bruno Lepri. 2017. Personality traits and ego-network dynamics. *PLOS ONE* 12, 3 (2017), e0173110.

[22] Rhea M Chase and Donna B Pincus. 2011. Sleep-related problems in children and adolescents with anxiety disorders. *Behavioral Sleep Medicine* 9, 4 (2011), 224–236.

[23] Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 16, 1 (1990), 22–29.

[24] Glen Coppersmith, Mark Dredze, and Craig Harman. 2014. Quantifying mental health signals in Twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. 51–60.

[25] Glen Coppersmith, Mark Dredze, Craig Harman, and Kristy Hollingshead. From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality* (2015). Association for Computational Linguistics, 1–10.

[26] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013a. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, 3267.

[27] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013b. Predicting Depression via Social Media. *ICWSM* 13 (2013), 1–10.

[28] Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, 2098–2110.

[29] David C DeAndrea, Nicole B Ellison, Robert LaRose, Charles Steinfield, and Andrew Fiore. 2012. Serious social media: On the use of social media for improving students' adjustment to college. *The Internet and higher education* 15, 1 (2012), 15–23.

[30] Afsaneh Doryab, Daniella K Villalba, Prerna Chikersal, Janine M Dutcher, Michael Tumminia, Xinwen Liu, Sheldon Cohen, Kasey Creswell, Jennifer Mankoff, John D Creswell, and Anind K Dey. 2019. Identifying Behavioral Phenotypes of Loneliness and Social Isolation with Passive Sensing: Statistical Analysis, Data Mining and Machine Learning of Smartphone and Fitbit Data. *JMIR mHealth and uHealth* 7, 7 (July 2019), e13209.

[31] Alexandra Ehrenberg, Suzanna Juckes, Katherine M. White, and Shari P. Walsh. 2008. Personality and Self-Esteem as Predictors of Young People's Technology Use. *CyberPsychology & Behavior* 11, 6 (2008), 739–741.

[32] Johannes C. Eichstaedt, Robert J. Smith, Raina M. Merchant, Lyle H. Ungar, Patrick Crutchley, Daniel Preoţiuc-Pietro, David A. Asch, and H. Andrew Schwartz. 2018. Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences* 115, 44 (Oct. 2018), 11203–11208.

[33] Jon D. Elhai, Mojisola F. Tiamiyu, Justin W. Weeks, Jason C. Levine, Kristina J. Picard, and Brian J. Hall. 2018. Depression and emotion regulation predict objective smartphone use measured over one week. *Personality and Individual Differences* 133 (Oct. 2018), 21–28.

[34] Nicole B. Ellison, Charles Steinfield, and Cliff Lampe. 2007. The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication* 12, 4 (July 2007), 1143–1168.

[35] Nicole B Ellison, Charles Steinfield, and Cliff Lampe. 2011. Connection strategies: Social capital implications of Facebook-enabled communication practices. *New Media & Society* 13, 6 (2011), 873–892.

[36] Norman S Endler and Nancy L Kocovski. 2001. State and trait anxiety revisited. *Journal of anxiety disorders* 15, 3 (2001), 231–245.

[37] Golnoosh Farnadi, Susana Zoghbi, Marie-Francine Moens, and Martine De Cock. 2013. Recognising personality traits using Facebook status updates. In *Seventh International AAAI Conference on Weblogs and Social Media*.

[38] Gillian Fergie, Shona Hilton, and Kate Hunt. 2016. Young adults' experiences of seeking online information about diabetes and mental health in the age of social media. *Health Expectations* 19, 6 (2016), 1324–1335.

[39] Katya C. Fernandez, Cheri A. Levinson, and Thomas L. Rodebaugh. 2012. Profiling: Predicting Social Anxiety From Facebook Profiles. *Social Psychological and Personality Science* 3, 6 (Nov. 2012), 706–713.

[40] Wendi L Gardner, Cynthia L Pickett, Valerie Jefferis, and Megan Knowles. 2005. On the outside looking in: Loneliness and social monitoring. *Personality and Social Psychology Bulletin* 31, 11 (2005), 1549–1560.

[41] Reshmi Gopalakrishna Pillai, Mike Thelwall, and Constantin Orasan. 2018. Detection of Stress and Relaxation Magnitudes for Tweets. In *Companion of the The Web Conference 2018 on The Web Conference 2018*. International World Wide Web Conferences Steering Committee, 1677–1684.

[42] Samuel D Gosling, Peter J Rentfrow, and William B Swann. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality* 37, 6 (Dec. 2003), 504–528.

[43] Rachel Grieve, Michaelle Indian, Kate Witteveen, G Anne Tolan, and Jessica Marrington. 2013. Face-to-face or Facebook: Can social connectedness be derived online? *Computers in Human Behavior* 29, 3 (2013), 604–609.

[44] Tom Griffiths. 2002. Gibbs sampling in the generative model of latent dirichlet allocation. (2002).

[45] Rebecca A Harris, Pamela Qualter, and Sarita J Robinson. 2013. Loneliness trajectories from middle childhood to pre-adolescence: impact on perceived health and sleep disturbance. *Journal of Adolescence* 36, 6 (2013), 1295–1304.

[46] Louise C Hawkley and John T Cacioppo. 2010. Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine* 40, 2 (2010), 218–227.

[47] Ron D Hays and M Robin DiMatteo. 1987. A short-form measure of loneliness. *Journal of Personality Assessment* 51, 1

(1987), 69–81.
[48] Thomas Holtgraves. 2011. Text messaging, personality, and the social context. *Journal of Research in Personality* 45, 1 (Feb. 2011), 92–99.
[49] Lindsay T Hoyt, Katharine H Zeiders, Natasha Chaku, Russell B Toomey, and Rajni L Nair. 2018. Young adults' psychological and physiological reactions to the 2016 US presidential election. *Psychoneuroendocrinology* 92 (2018), 162–169.
[50] Michaelle Indian and Rachel Grieve. 2014-03. When Facebook is easier than face-to-face: Social support derived from Facebook in socially anxious individuals. *Personality and Individual Differences* 59 (2014-03), 102–106.
[51] Borae Jin. 2013. How lonely people use and perceive Facebook. *Computers in Human Behavior* 29, 6 (Nov. 2013), 2463–2470.
[52] Tuomas Kari, Jonne Arjoranta, and Markus Salo. 2017. Behavior change types with Pokémon GO. In *Proceedings of the International Conference on the Foundations of Digital Games - FDG '17*. ACM Press, Hyannis, Massachusetts, 1–10.
[53] Ethan Kross, Philippe Verduyn, Emre Demiralp, Jiyoung Park, David Seungjae Lee, Natalie Lin, Holly Shablack, John Jonides, and Oscar Ybarra. 2013. Facebook use predicts declines in subjective well-being in young adults. *PLOS ONE* 8, 8 (2013), e69841.
[54] Dekang Lin. 1998. Extracting collocations from text corpora. In *First workshop on computational terminology*. Citeseer, 57–63.
[55] Shuai-Ting Lin, Pinchen Yang, Chien-Yu Lai, Yu-Yun Su, Yi-Chun Yeh, Mei-Feng Huang, and Cheng-Chung Chen. 2011. Mental Health Implications of Music: Insight from Neuroscientific and Clinical Studies. *Harvard Review of Psychiatry* 19, 1 (2011), 34–46.
[56] Edward Loper and Steven Bird. 2002. NLTK: the natural language toolkit. *arXiv preprint cs/0205028* (2002).
[57] Kate Loveys, Patrick Crutchley, Emily Wyatt, and Glen Coppersmith. 2017. Small but Mighty: Affective Micropatterns for Quantifying Mental Health from Social Media Language. In *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality*. Association for Computational Linguistics, 85–95. http://aclweb.org/anthology/W17-3110
[58] Andrew Kachites McCallum. 2002. MALLET: A Machine Learning for Language Toolkit. (2002). http://mallet.cs.umass.edu.
[59] David C. Mohr, Mi Zhang, and Stephen M. Schueller. 2017. Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. *Annual Review of Clinical Psychology* 13, 1 (May 2017), 23–47.
[60] Raheel Mushtaq, Sheikh Shoib, Tabindah Shah, and Sahil Mushtaq. 2014. Relationship between loneliness, psychiatric disorders and physical health ? A review on the psychological aspects of loneliness. *Journal of Clinical and Diagnostic Research : JCDR* 8, 9 (Sept. 2014).
[61] Robin L. Nabi, Abby Prestin, and Jiyeon So. 2013. Facebook Friends with (Health) Benefits? Exploring Social Network Site Use and Perceptions of Social Support, Stress, and Well-Being. *Cyberpsychology, Behavior, and Social Networking* 16, 10 (Oct. 2013), 721–727.
[62] John A. Naslund, Stuart W. Grande, Kelly A. Aschbrenner, and Glyn Elwyn. 2014. Naturally Occurring Peer Support through Social Media: The Experiences of Individuals with Severe Mental Illness Using YouTube. 9, 10 (2014), e110171.
[63] Dag Neckelmann, Arnstein Mykletun, and Alv A. Dahl. 2007. Chronic Insomnia as a Risk Factor for Developing Anxiety and Depression. *Sleep* 30, 7 (2007), 873–880.
[64] NIMH. 2020. Depression. (2020). https://www.nimh.nih.gov/health/topics/depression/index.shtml
[65] Alicia L Nobles, Jeffrey J Glenn, Kamran Kowsari, Bethany A Teachman, and Laura E Barnes. 2018. Identification of imminent suicide risk among young adults using text messages. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 413.
[66] Jiyoung Park, David Seungjae Lee, Holly Shablack, Philippe Verduyn, Patricia Deldin, Oscar Ybarra, John Jonides, and Ethan Kross. 2016. When perceptions defy reality: the relationships between depression and actual and perceived Facebook social support. *Journal of Affective Disorders* 200 (2016), 37–44.
[67] Sungkyu Park, Inyeop Kim, Sang Won Lee, Jaehyun Yoo, Bumseok Jeong, and Meeyoung Cha. 2015. Manifestation of Depression and Loneliness on Social Networks: A Case Study of Young Adults on Facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. ACM Press, 557–570.
[68] Sachin R Pendse. 2018. Sochiatrist: Inferring the Relationship Between Emotion and Private Social Messages. (2018).
[69] Cynthia L. Pickett and Wendi L. Gardner. 2005. The Social Monitoring System: Enhanced Sensitivity to Social Cues as an Adaptive Response to Social Exclusion. *The social outcast: Ostracism, social exclusion, rejection, and bullying.* (2005), 213–226. Place: New York, NY, US Publisher: Psychology Press.
[70] Cynthia L Pickett, Wendi L Gardner, and Megan Knowles. 2004. Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality and Social Psychology Bulletin* 30, 9 (2004), 1095–1107.
[71] Donna J. Reid and Fraser J.M. Reid. 2007. Text or Talk? Social Anxiety, Loneliness, and Divergent Preferences for Cell Phone Use. *CyberPsychology & Behavior* 10, 3 (June 2007), 424–435.

[72] Andreas Reissmann, Joachim Hauser, Ewelina Stollberg, Ivo Kaunzinger, and Klaus W Lange. 2018. The role of loneliness in emerging adults' everyday use of Facebook–an experience sampling approach. *Computers in Human Behavior* 88 (2018), 47–60.

[73] Eric Rice, Seth Kurzban, and Diana Ray. 2012. Homeless But Connected: The Role of Heterogeneous Social Network Ties and Social Networking Technology in the Mental Health Outcomes of Street-Living Adolescents. 48, 6 (2012), 692–698.

[74] Laura Smart Richman, Julie Martin, and Jennifer Guadagno. 2016. Stigma-based rejection and the detection of signs of acceptance. *Social Psychological and Personality Science* 7, 1 (2016), 53–60.

[75] Dan Russell, Letitia A Peplau, and Carolyn E Cutrona. 1980. The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. *Journal of personality and social psychology* 39, 3 (1980), 472.

[76] Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, and David C Mohr. 2015. Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. *Journal of Medical Internet Research* 17, 7 (July 2015), e175.

[77] Koustuv Saha, Ayse E Bayraktaroglu, Andrew T Campbell, Nitesh V Chawla, Munmun De Choudhury, Sidney K D'Mello, Anind K Dey, Ge Gao, Julie M Gregg, Krithika Jagannath, and others. 2019. Social media as a passive sensor in longitudinal studies of human behavior and wellbeing. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–8.

[78] Akane Sano, Andrew J Phillips, Z Yu Amy, Andrew W McHill, Sara Taylor, Natasha Jaques, Charles A Czeisler, Elizabeth B Klerman, and Rosalind W Picard. 2015. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–6.

[79] H. Andrew Schwartz, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin E. P. Seligman, and Lyle H. Ungar. 2013a. Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE* 8, 9 (Sept. 2013), e73791.

[80] H. Andrew Schwartz, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin E. P. Seligman, and Lyle H. Ungar. 2013b. Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE* 8, 9 (09 2013), 1–16.

[81] H. Andrew Schwartz, Maarten Sap, Margaret L. Kern, Johannes C. Eichstaedt, Adam Kapelner, Megha Agrawal, Eduardo Blanco, Lukasz Dziurzynski, Gregory Park, David Stillwell, Michal Kosinski, Martin E.P. Seligman, and Lyle H. Ungar. 2016. Predicting Individual Well-Being Through The Language of Social Media. In *Biocomputing 2016*. World Scientific, Kohala Coast, Hawaii, USA, 516–527.

[82] Kevin S. Seybold and Peter C. Hill. 2001. The Role of Religion and Spirituality in Mental and Physical Health. *Current Directions in Psychological Science* 10, 1 (2001), 21–24.

[83] Phillip Shaver, Wyndol Furman, and Duane Buhrmester. 1985. Transition to college: Network changes, social skills, and loneliness. (1985).

[84] Ariel Shensa, Jaime E Sidani, César G Escobar-Viera, Kar-Hai Chu, Nicholas D Bowman, Jennifer M Knight, and Brian A Primack. 2018. Real-life closeness of social media contacts and depressive symptoms among university students. *Journal of American College Health* 66, 8 (2018), 747–753.

[85] Marcin Skowron, Marko Tkalčič, Bruce Ferwerda, and Markus Schedl. 2016. Fusing social media cues: personality prediction from twitter and instagram. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 107–108.

[86] C. D. Spielberger, R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs. 1983. *Manual for the State-Trait Anxiety Inventory.* Consulting Psychologists Press.

[87] Jacopo Staiano, Bruno Lepri, Nadav Aharony, Fabio Pianesi, Nicu Sebe, and Alex Pentland. 2012. Friends don't lie: inferring personality traits from social network structure. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 321–330.

[88] Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29, 1 (2010), 24–54.

[89] Svitlana Volkova, Kyungsik Han, and Courtney Corley. 2016. Using social media to measure student wellbeing: a large-scale study of emotional response in academic discourse. In *International Conference on Social Informatics*. Springer, 510–526.

[90] Tao Wang, Markus Brede, Antonella Ianni, and Emmanouil Mentzakis. 2017. Detecting and characterizing eating-disorder communities on social media. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM, 91–100.

[91] Xinyu Wang, Chunhong Zhang, and Li Sun. 2013. An improved model for depression detection in micro-blog social network. In *2013 IEEE 13th International Conference on Data Mining Workshops*. IEEE, 80–87.

[92] Jason Wilson. 2016. Pokémon Go launches in US on iOS and Android. (2016). https://venturebeat.com/2016/07/06/pokemon-go-launches-worldwide-on-ios-and-android/
[93] David H. Wolpert. 1992. Stacked Generalization. *Neural Networks* 5 (1992), 241–259.
[94] Kevin B. Wright. 2012. Emotional Support and Perceived Stress Among College Students Using Facebook.com: An Exploration of the Relationship Between Source Perceptions and Emotional Support. *Communication Research Reports* 29, 3 (July 2012), 175–184.
[95] Congyu Wu, Mehdi Boukhechba, Lihua Cai, Laura E Barnes, and Matthew S Gerber. 2018. Improving momentary stress measurement and prediction with bluetooth encounter networks. *Smart Health* (2018).
[96] Chia-chen Yang and Yen Lee. 2018. Interactants and activities on Facebook, Instagram, and Twitter: associations between social media use and social adjustment to college. *Applied Developmental Science* (2018), 1–17.
[97] Marc Ziegele and Leonard Reinecke. 2017. No place for negative emotions? The effects of message valence, communication channel, and social distance on users' willingness to respond to SNS status updates. *Computers in Human Behavior* 75 (2017), 704 – 713.

Received January 2020; revised June 2020; accepted July 2020