Detecting and Characterizing Eating-Disorder Communities on Social Media

Tao Wang
ESRC DTC, IfLS

Markus Brede
Department of ECS

Antonella Ianni
Department of Economics

Emmanouil Mentzakis
Department of Economics

University of Southampton, UK
{t.wang, Markus.Brede, A.Ianni, E.Mentzakis}@soton.ac.uk

ABSTRACT

Eating disorders are complex mental disorders and responsible for the highest mortality rate among mental illnesses. Recent studies reveal that user-generated content on social media provides useful information in understanding these disorders. Most previous studies focus on studying communities of people who discuss eating disorders on social media, while few studies have explored community structures and interactions among individuals who suffer from this disease over social media. In this paper, we first develop a snowball sampling method to automatically gather individuals who self-identify as eating disordered in their profile descriptions, as well as their social network connections with one another on Twitter. Then, we verify the effectiveness of our sampling method by: 1. quantifying differences between the sampled eating disordered users and two sets of reference data collected for non-disordered users in social status, behavioral patterns and psychometric properties; 2. building predictive models to classify eating disordered and non-disordered users. Finally, leveraging the data of social connections between eating disordered individuals on Twitter, we present the first homophily study among eating-disorder communities on social media. Our findings shed new light on how an eating-disorder community develops on social media.

1. INTRODUCTION

Eating disorders (ED) are mental disorders characterized by abnormal attitudes towards food and unusual eating habits [4, 36]. The most common ED are anorexia nervosa where sufferers restrict their eating to keep low weight, and bulimia nervosa where people repeat cycles of binge eating and purging [4]. ED can negatively affect a person's physical, psychological and social health. Symptoms of ED include self-starvation, laxative abuse, anxiety, depression, or other extreme behavioral and emotional responses to eating food and gaining weight [36]. Specifically, ED exhibit the highest mortality rate of any mental illness and 20% of all deaths from anorexia are the result of suicide [2]. Recently, preva-

lence of ED has significantly grown, with a 2015 report from the ED charity Beat estimating more than 725,000 people in the UK are eating disordered (ED-ed), in a trend that is increasing over time [7]. More than 85% of those suffering are below the age of 19 and 95% of sufferers are females.

To date, numerous studies in psychiatry, psychology and medicine have been carried out to identify and understand ED [22, 40, 1]. However, these clinical studies are typically carried out by means of surveys and interviews, and these methods are known to present several limitations. First, the denial of illness, ambivalence towards treatment and high drop-out rates make ED-ed populations hard to detect and reach [24]. Even in cases where data can be collected, participants may conceal their condition and/or its extent, largely reducing the response accuracy and reliability of the data. Second, most of these surveys and interviews are conducted within small groups of individuals in a temporal granularity, which may not be representative of large populations and miss finer-grained longitudinal data [20]. Finally, pre-defined questionnaires alone may be insufficient to reveal the physical and psychological states of individuals.

The usage of social media (SM) services, such as Twitter, Facebook and Instagram, to express and exchange thoughts or to document details of daily life has increased steadily over recent years, particularly in young populations. Previous studies have shown that people's behaviors and content generated on SM can indeed be used to infer their mental health states [26, 14, 10]. In this paper, we show that using analyses based on SM data can help to overcome the limitations of traditional surveys in ED studies, by providing finer-grained features of ED with a large number of samples. Besides, the (semi-)anonymous nature of SM encourages people to naturally socialize and self-disclose [6], and this allows us to study ED by utilizing naturally occurring data in a non-reactive way. Thus, SM data may complement conventional data and help estimating the offline occurrences of ED.

Some computational methods have been proposed to study ED and other mental illnesses over SM recently [18, 14, 25, 19, 16, 10, 11, 12]. Most previous studies focus on identifying signs of a mental illness from user-generated content on SM. However, few studies have explored social ties and interactions between mentally ill peers over SM. A rich body of psychological literature shows that people's concerns and behaviors can be influenced by peer pressure [37]. Evidence suggests that the social dimension captured by social networks plays an important role in the studies of life-style related conditions, such as ED and other mental disorders [27]. In this paper, we explore an alternative and complementary

method to detect ED-ed communities and characterize social interactions among ED-ed peers on SM, focusing on Twitter. The main contributions of this work are:

- We present a snowball sampling method to sift ED-ed individuals and their social networks from Twitter. Unlike prior methods that gather data by surveys or by filtering users who post ED-related content [18, 46, 11], we sample individuals who self-identify as ED-ed in their profile descriptions on Twitter and expand the sample group with snowball sampling through their social networks of followees/followers, thereby recovering connected communities of individuals who are likely to display ED on Twitter.
- Comparing the differences between ED-ed and two sets of non-ED-ed users in social status, behavioral patterns and psychometric properties, we show that our sampled ED dataset captures key characteristics of ED, e.g., young ages, prevailing urges to lose weight even if being clinically underweight, high social anxiety, intensive self-focused attention, deep negative emotion, increased mental instability, and excessive concerns of body image and ingestion.
- We show that users' behaviors and content generated on Twitter can help to identify whether or not a user is affected by ED by training SVM classifiers to distinguish between ED-ed and non-ED-ed users. The classifiers have achieved an accuracy of more than 97%, and the differences of ED-ed and non-ED-ed users are more easily distinguishable than those between two sets of non-ED-ed users. This further confirms the reliability of our sampling method in targeting ED-ed populations on Twitter.
- Using the social networking data between ED-ed users, we investigate the social interactions among ED-ed peers and explore the presence of homophily in the ED-ed communities on Twitter. We find that ED-ed users who show similar tweeting preferences, concerns about death, habits in using language and body weight tend to preferentially interact with one another. These findings can help to understand the way an ED-ed community develops on SM.

Our empirical findings shed light on developing automated techniques to sample data for larger communities with ED or other mental disorders from SM, beyond those that have self-identified. This work also lays the basis for analytically characterizing ED-ed communities and their social structures on SM, thereby providing guidance to develop effective interventions not only for individuals but for large groups.

2. RELATED WORK

Most previous studies of ED based on SM data come from the psychological and medical communities. Juarascio et al. [26] make personal contacts with several pro-ED groups on Facebook and MySpace to get access to observe and analyze the groups' content. Wolf et al. [45] analyze pro-ED, recovery and control blogs relying on quantitative text analyses. Syed-Abdul et al. [41] study anorexia-related misinformation propagated through YouTube videos. Arseniev-Koehler et al. [3] find that many followers of pro-ED users also self-identify with ED by studying 45 pro-ED users on Twitter. Most of these studies use qualitative methods and involve intensive manual labor in data collection and validation. This work contributes to this literature by developing computational techniques to automatically detect and quantitatively analyze ED-ed communities on SM platforms.

In the social computing community, research on ED over SM is limited, especially on Twitter. Twitter, which was created in 2006 and is used by 33% of US teens [29], provides *rich* and *public* information of users' social and behavioral context. Analyzing such information can offer a deep insight into ED-ed individuals. Moreover, while many sites such as Facebook and Instagram have taken steps to counteract the diffusion of pro-ED content [12], Twitter has taken no actions to limit such content [3]. The latter feature makes Twitter a unique online SM platform to study ED. A recent work measures the psychological features in a "pro-anorexia" community on Twitter [46]. However, the community studied in this work is a group of users who talk about ED in their tweets, and this typically includes not only people who are really affected by the condition, but also a large number of people who casually discuss the disease on a one-off basis. By contrast, in what follows, we study the community of people who self-identify with ED in their profile descriptions, and we show that such information is more reliable.

Research on ED has also been carried out on SM platforms other than Twitter. The differences of pro-anorexia and pro-recovery posts on Tumblr are studied in [16]. Another work further explores to predict the likelihood of a user in the recovery from ED on Tumblr [11]. Very recently, researchers study ED from the content of tags on Instagram: [10] quantifies the severity of ED for a collection of users who post ED-related tags; [12] examines the content moderation and lexical variation in ED-related users; [9] measures the characteristics of removed content about ED. While the findings in these studies are insightful, they are mostly confined to the study on individuals' behavioral patterns in isolation, without their mutual interactions. This work extends prior work by studying social interactions in ED-ed groups on SM. The work of [47] is the most closely related to ours, as the authors examine the interactions between pro-anorexia and pro-recovery communities on Flickr. We extend this research by exploring individuals' attributes that can facilitate the social interactions in ED-ed communities on SM.

3. DATA

All data we gathered is *public* information on Twitter, and available via the official Twitter API. Any data that has been set as *private* is not included in our study.

3.1 Collecting ED Data

A big challenge faced by research on ED and on other mental illnesses from SM is how to gather a sufficient number of reliable sample individuals with an illness (i.e., positive samples). To date, researchers seek positive samples mainly relying on users' self-reported diagnoses [18, 14, 16, 10]. The methods of collecting self-reports are broadly classified into two categories. One category is *survey* based methods, in which self-reports are gathered by surveys (via personal contacts or crowd-sourcing) [26, 18, 19]. Typically, survey based methods are time-consuming and expensive to create a large sample set, and often suffer from small sample sizes. This feature may in fact undermine the statistical significance of the results obtained by these methods. The other category is *information filtering* based methods, in which self-reports are filtered from public information available online by using computational techniques [14, 25, 16, 10, 9]. Most filtering methods use a set of keywords as search queries to filter users whose posts on SM (e.g., tweets on Twitter) contain

these keywords. Due to pervasive noise in online information, these methods often suffer from low quality of data. Moreover, existing data collection methods mainly focus on gathering positive individuals, but missing the data of social network connections between individuals.

3.1.1 Filtering Self-reported ED-ed Diagnoses

To retrieve reliable ED-ed samples, we draw self-reported diagnoses from users' profile descriptions on Twitter, i.e., the user-defined texts describing their accounts below profile images. This is based on two observations. First, a profile description is often regarded as the biography of a user, while many statements in tweets are less trustworthy. Second, users' personal profiles always pertain to themselves, while people often comment on (or refer to) others in tweets. Thus, the information in profiles may be more indicative of the most genuine aspect of a user than that in tweets [15]. Table 1 shows some examples of diagnostic statements in tweets and profile descriptions. We see that people who talk about ED in tweets may not be affected by ED. It may be difficult to identify a user as ED positive based on one of their posts.

Table 1: Examples of self-reported ED-ed diagnoses.

Diagnostic Statements in Tweets	
Joke	My mom and brother thinks I have a **eating disorder** cause I don't eat a lot and when I do eat my mom tells me 'good girl'.
Reference	If you're saying @USER has an **eating disorder**, please unfollow me.
Diagnostic Statements in Profile Descriptions	
ED-related User	Project HEAL Toronto Chapter! Promoting **Eating Disorder** Awareness, Positive Body Image, & Scholarships for Those Battling ED.
ED-ed User	16 years old. **Ednos**. Not skinny enough for **anorexia**. 128 **lbs** 5 foot 7 inch / Yes i look and feel like a wale. I will reach my **UGW**; 99 **lbs**.

Based on the above considerations, we assume that users self-identify as being diagnosed with ED, for the purposes of our study, if their profile descriptions display any ED-diagnosis keywords listed in Table 2. These keywords are initialized with the semantically related words of "eating disorder" in the Urban Dictionary[1]. Urban Dictionary is a crowd-sourced online dictionary of slang words and phrases, which is useful to find the words that are currently popular on the Internet. Then, we finalize the ED-diagnosis keywords with the following processing.

1. Remove words that are generic to use in various non-ED contexts, such as "food", "fat" and "self-harm".
2. Remove words that are the abbreviations of ED and ED-related symptoms but have ambiguity, such as "ed" (may denote the past-tense suffix of verbs), "ana" and "mia" (may be a person's name).
3. Track the tweet stream with the keywords after the above refinements via the Twitter Public API, and add words, which have high co-occurrences with the refined keywords in the crawled tweets and can directly map to ED, into the ED-diagnosis keyword set.

However, users whose profile descriptions display ED-related keywords may be therapists, institutes or other organizations related to ED rather than genuine ED-ed sufferers (see

Table 1 for example). To filter out these ED-related but non-ED-ed users and obtain higher-quality samples, we further add another filtering constraint by requiring that users' profile descriptions should also contain some personal biological information (bio-information), such as body weight. As organizations have no such bio-information and ED therapists are unlikely to disclose bio-information in their profiles, this constraint can help to refine our ED-ed samples. Since ED-ed sufferers generally focus excessively on their body weight [4, 36], most bio-information keywords we used are weight-related words. Table 2 lists the bio-information keywords we used and their descriptions. Some of these words are the acronyms of ED glossary[2] and the remainder are the units of weight (e.g., lbs and kg). Profile descriptions are considered to disclose bio-information if they contain any of these bio-information keywords. We identify a user as ED positive if their Twitter profile descriptions contain both ED-diagnosis information and personal bio-information.

Table 2: Keywords of ED-ed diagnoses and bio-information used for filtering ED-ed users.

Category	Keywords
ED diagnosis	eating disorder, eatingdisorder, anorexia, anorexic, anorexia nervosa, bulimia, bulimic, bulemia, bulimia nervosa, ednos, edprob, proana, promia, anamia, askanamia, purge, binge, thinspo, bonespo, legspo.
Bio-information	BMI (Body Mass Index), CW (Current Weight), UGW (Ultimate Goal Weight), GW (Goal Weight), HW (Highest Weight), LW (Lowest Weight), lbs, kg.

3.1.2 Snowball Sampling ED-ed Communities

To obtain a larger number of ED-ed users and their social connections, we develop a user collection method based on snowball sampling. Algorithm 1 shows the detailed steps of this method. The sampling is carried out via breadth-first search. Line 1 to line 5 show the initialization of seed users. Line 6 to line 13 show the snowball sampling to collect ED-ed communities. $Publishers(T)$ denotes the set of unique users who published the set of initial tweets T. Function $ED\_check(u_i, K_d, K_b)$ returns true if the profile description of user u_i contains at least one ED-diagnosis keyword in K_d and one bio-information keyword in K_b. $V^{(l)}$ denotes the subset of users sampled at level l. $Friends(u_i)$ denotes u_i's friends on Twitter, including followers and followees. In the updates of edges E (line 12), add $e(u_i, u_j)$ if u_j is one of u_i's followers, and add $e(u_j, u_i)$ if u_j is one of u_i's followees. Our crawler implemented based on Algorithm 1 stops after six rounds of snowballing in February 2016. At each sampling stage, we filter out non-English speaking accounts and finally obtain 3,380 unique users.

To inspect the quality of our collected ED data, we develop a labeling system by which the Twitter homepage of each user is automatically downloaded for inspectors. Inspectors annotate each user as to whether a user is suspected of having ED according to their posted tweets, images and friends' profiles. Our annotation results on randomly selected 1,000 samples show that almost all of the checked samples are suspected of having ED and 95.2% of the samples are labeled as being highly likely to have ED. This illustrates that the proposed data collection method provides

---

[1]http://www.urbandictionary.com, retrieved January 2016.

[2]http://glossary.feast-ed.org/

Algorithm 1: Snowball sampling framework for collecting ED-ed communities on Twitter.

Input: Community graph $G = (V, E)$ with user set V and directed follow edge set E; set of ED-diagnosis keywords K_d; set of bio-information keywords K_b.
Output: G.

1 $(V, E) \leftarrow (\emptyset, \emptyset)$;
2 Track initial tweet stream T with K_d;
3 **for** $u_i \in Publishers(T)$ **do**
4   **if** $ED\_check(u_i, K_d, K_b)$ **then**
5       $V^{(0)} \leftarrow V^{(0)} + \{u_i\}$;
6 $l \leftarrow 0$;
7 **while** $V^{(l)} \neq \emptyset$ **do**
8   **for** $u_i \in V_l$ **do**
9   **for** $u_j \in Friends(u_i)$ **do**
10   **if** $ED\_check(u_j, K_d, K_b)$ **and** $u_j \notin V$ **then**
11             $V^{(l+1)} \leftarrow V^{(l+1)} + \{u_j\}$;
12             $E \leftarrow E + \{e(u_i, u_j) | e(u_j, u_i)\}$;
13    $l \leftarrow l + 1$;
14 **return** G;

a set of relatively high-quality ED positive samples. Moreover, in the following section, we further use classifiers to verify the reliability of our data collection method.

All the 3,380 users are used as ED-positive samples. For each ED-ed user, we download up to 3,200 (the limit returned from Twitter official API) of their most recent tweets.

3.2 Collecting Reference Data

To validate our sampled ED data, we collect two sets of reference data as negative samples. The first set of data is used to compare the differences between ED-ed users and the general population on Twitter, which is built by selecting a set of users at random, labeled as *Random* data. The second set of data is used to compare the differences between ED-ed users and young females, labeled as *Younger* data. As ED develop predominantly in young females [1, 4], the effects of demographics (i.e., age and gender) can be further controlled in comparing ED and Younger users, which helps to explore the key differences between ED-ed and non-ED-ed users.

Random Data. We construct Random data as follows. First, 252,970 initial tweets are randomly sampled via the Twitter Public API. To avoid biases of sampling tweets about specific topics or from specific communities, we collect these tweets in three phases over two weeks. In each phase, only tweets written in English are collected. Second, from the unique users who posted these initial tweets, 3,380 (the same number of ED-ed users) users are randomly selected. Third, to avoid another bias of preferentially sampling users that are very active on Twitter, we further crawl the friends (including followees and followers) of the 3,380 seed users. Finally, we randomly select 30,684 users and retrieve their most recent tweets to finalize Random data.

Younger Data. To target young female populations on Twitter, we use the names of 14 popular artists, ranked by Billboard[3] in 2016, as keywords to track an initial tweet stream of candidate users. This is motivated by the observation that popular music is always a hot topic discussed among young people on Twitter. The initial tweets are also filtered in three phases. Then, we refine the candidate users

[3]http://www.billboard.com/artists/top-100/2016

by filtering female users. To this end, we follow the widely used method in previous work [38, 17], i.e., 1. select the candidate users that have given a full name in their profiles; 2. perform a lexicon-based method that identifies matches of the first name of each selected user to a dictionary of first names. Our name dictionary is built with the top 200 most popular first names for girls born in 2000s, obtained from the US Social Security Administration[4]. Also, we filter out the accounts that have been verified to exclude celebrity friends of the listed artists. Next, we select 3,380 refined users as seed users and crawl their friends (only the non-verified users with a female name are collected). Finally, we randomly select 37,983 users and download their most recent tweets to finalize Younger data.

Table 3: Statistics of numbers of users, numbers of tweets, and average numbers of tweets per user.

Dataset	#Users	#Tweets	#T/U
ED	3,380	1,797,239	531.73
Random	30,684	60,774,175	1,980.65
Younger	37,983	57,253,947	1,507.36

Table 3 lists the statistics on the three sets of data. There are no pairwise intersections between all of the user sets.

4. USER CHARACTERIZATION

4.1 Measures

We first present three types of measures to characterize differences between ED-ed and non-ED-ed users on Twitter.

4.1.1 Social Status

Engagement. We define three engagement measures based on the overall volumes of users' followees, tweets and followers respectively, to assess users' states of being engaged on Twitter. However, previous studies report that many statistics of users on Twitter obey power-law distributions [34, 30]; some Twitter users have posts and social connections that greatly exceed the average. To reduce the skewness towards large values, we employ logarithmic scales and define the engagement degree of user u in terms of statistic s as:

$$Engagement(u, s) = \log(1 + \#s_u), \quad (1)$$

where $s \in \{Followees, Tweets, Followers\}$, and $\#s_u$ denotes the count of s that u has. The constant 1 is added to avoid infinite values in logarithmic scales.

Activity. Similarly, we use the average normalized numbers of followees, tweets and followers per day to measure the activity of a user on Twitter. The activity degree of user u in terms of statistics s is defined as:

$$Activity(u, s) = \log\left(1 + \frac{\#s_u}{t_u}\right), \quad (2)$$

where t_u denotes the number of days from the date of u joining Twitter to the date of u's last post.

4.1.2 Behavioral Patterns

Tweeting Preference. We use the proportions of tweets that involve different types of behaviors in a user' most recent tweets to measure users' tweeting preferences. The behaviors of interest are: three manners of publishing posts

[4]https://www.ssa.gov/oact/babynames/decades/names2000s.html

(i.e., originally tweet, re-tweet and quote[5]); two forms of interacting with others (i.e., mention and reply); following the discussions of public topics (i.e., use hashtags); and sharing external links (i.e., append URLs). Note that we only count the mentions that are directly made by users. Any mentions in the original tweets that users re-tweeted are ignored.

Interaction Diversity. We also quantify the ways in which users interact with the external world, specifically on examining whether a user tends to follow a variety of topics or a specific set of topics; whether she/he prefers to interact with various individuals or certain specific individuals. For this purpose, we employ entropy, which is widely used in previous studies [21, 44], as a diversity measure. Given a user u, we track the sequence of targets of interest to u (e.g., hashtags u used or other users u re-tweeted in the past), denoted as T_u. The interest diversity of u in terms of a type of interactions I is computed by calculating the entropy of such interactions with different targets $v \in T_u$:

$$H(u, I) = - \sum_{v \in T_u} p(I_v) \log p(I_v), \qquad (3)$$

where $I \in \{Hashtag, Re\text{-}tweet, Mention, Reply\}$, and $p(I_v) = \frac{\#I_v}{\sum_{j \in T_u} \#I_j}$. $\#I_v$ is the number of interactions I with target v, e.g., using hashtag v or re-tweeting user v. Larger entropy values indicate a higher diversity of interests that a user has.

4.1.3 Psychometric Properties

We use the Linguistic Inquiry and Word Count (LIWC) lexicon [42] to distill a set of variables that relate to health statistics from users' posts on Twitter. LIWC is composed of 80 psychologically-relevant categories and about 4,500 word patterns[6]. Each word pattern is associated with one or more categories, corresponding to different emotions, linguistic styles, personal concerns, etc. Given a text file, LIWC computes the percentage of words that match each of these built-in categories, and hence produces a quantitative summary of 80 dimensions for the textual data. This lexicon has been widely used to capture people's psychological and health states from the words they use [15, 14, 19].

To facilitate this analysis, we combine the collection of posts generated by each user together as a document. Users' re-tweets are also used in this analysis, as the content users re-tweeted can indicate their interests as well [33]. Then, we remove mentions, hashtags, URLs and the prefixes of re-tweets (i.e., "RT"). Finally, the pruned documents are split into tokens by white-space characters and the documents that have more than 50 tokens are processed with LIWC, resulting in a vector of category percentages for each user.

4.2 Classification Framework

Next, we follow prior studies [14, 25] that use classifiers to further verify the reliability of data sampling method. If the labels of ED-ed users are reliable, we expect that the performance of classifying ED-ed and non-ED-ed users would be better than that of classifying two sets of non-ED-ed users.

We build separate binary classifiers to predict the classes of ED, Random and Younger users. Each user is represented as a vector of 97 features obtained from the above measures (6 social-status features; 11 behavioral features; 80 psychometric features). To boost the performance of classification,

we standardize the values of each feature by subtracting the corresponding mean and dividing by the corresponding standard deviation. To determine the optimal classification algorithm, we compare several different parametric classifiers and non-parametric classifiers, such as Naive Bayes, Support Vector Machine (SVM) with linear, RBF, Sigmoid and Polynomial (degree=3) kernels and k-Nearest Neighbors with different settings on number of neighbors and distance functions, in our preliminary experiments. The best performing classifier we found is the linear SVM with the default settings in Scikit-learn 0.17 package[7]. As the samples are unbalanced across positive and negative classes, we adjust the regularization constants of different classes with the weights that are inversely proportional to class sizes in the training data [39]. To obtain more generalizable evaluations, all results are obtained with 5-fold cross validation. Each fold contains approximately the same percentage of users of each class.

5. COMMUNITY CHARACTERIZATION

Individuals connect/interact with others and form communities on Twitter primarily by four ways: "follow", "re-tweet", "reply" and "mention". According to follow, re-tweet, reply and mention ties between users (e.g., who-follows-whom ties), we build four types of weighted and directed networks among ED-ed users, i.e., follow, re-tweet, reply and mention networks, respectively. The re-tweet, reply and mention ties are extracted from users' most recent posts we retrieved. Typically, users establish different types of relational ties for different purposes, e.g., follow others to maintain a long-term friendship; re-tweet someone to diffuse information; mention and reply to a user for creating temporary conversations. Hence, different types of networks can have different topologies and reflect the features of a community from different perspectives [13]. Next, we investigate the characteristics of ED-ed communities based on these networks.

5.1 Network Characterization

We first examine the topological features of different types of networks built above. We measure networks by using nine widely used metrics: 1. total number of nodes (i.e., users); 2. total number of edges; 3. edge density (the ratio of number of edges to maximum possible number of edges); 4. average shortest path length of connected node pairs; 5. total number of weakly connected components; 6. fraction of nodes in the giant weak component; 7. global clustering coefficient (the probability that two neighbors of a node are connected); 8. reciprocity (the likelihood of nodes with mutual links); 9. assortativity coefficient of degree (the preference for nodes to link to others with similar degree values) [35]. Note that directed networks are considered as undirected ones in measuring global clustering coefficient, and the degree assortativity measured here are the correlations between source in-degree and destination out-degree [34].

5.2 Homophily Analysis

Homophily (known as assortative mixing in network science) is the tendency of individuals to connect with others who share similar characteristics [32]. The properties of homophily can help understand the way a community develops.

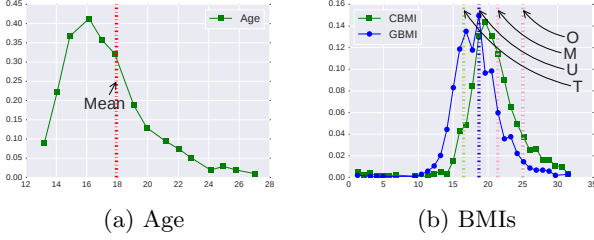---

[5]Add comments before re-tweeting to make it a quote tweet.
[6]The version used in this work is LIWC2007.

[7]http://scikit-learn.org/stable/index.html#

(a) Age   (b) BMIs

Figure 1: Probability density functions of ages and BMIs. "T", "U", "M" and "O" mark thinness, underweight, median and overweight cut-offs from WHO.

To explore homophily in ED-ed communities, we study assortative mixing in their networks built above, focusing on mixing according to social-status, behavioral and psychometric features measured above. We quantify the assortativity coefficients of different types of networks by each of these features [35]. To further test the statistical significance of these assortativity outcomes, for each feature, we randomly shuffle users' feature values and re-measure assortativity coefficients based on the shuffled values. We repeat this procedure 3,000 times to yield the simulated distributions of assortativity coefficients for each feature. Finally, we use two-tailed hypothesis tests to assess how significantly the assortativity outcomes differ from the simulated distributions.

6. RESULTS

6.1 ED Validation with Bio-information

Table 4: Statistics of bio-information in ED-ed users.

Category	Indicator	#Users	%Users
Observed Information	Age	1,030	30.47%
	Height	1,401	41.45%
	GW	2,781	82.28%
	CW	2,238	66.21%
	LW	466	13.79%
	HW	1,296	38.34%
Inferred Information	GBMI	1,168	34.56%
	CBMI	1,025	30.33%

We first validate our ED data from users' bio-information. We build several regular expressions to extract ED-ed users' bio-information, such as age, height, GW, CW, LW and HW (see Table 2 for notations), from their profile descriptions. According to the values of CW, GW and height, we further infer users' goal BMI (GBMI) and current BMI (CBMI) values respectively. Table 4 shows the number and percentage of users that have information related to each indicator. We see that the proposed data collection method has harvested a large amount of bio-information for ED-ed users.

Next, we discuss these indicators in detail. Figure 1 shows the distributions of age and BMIs values of ED-ed users. Consistent with the findings in clinical studies [1, 4], most targeted ED-ed users are teenage, concentrated in the age range of [14, 20], and the average age is 18. Comparing the curves of BMIs, we see that the GBMI values of this group of users are smaller than their CBMI values. This indicates that most of these users wish to lose weight, an important signal of ED [4]. Also, we obtain the reference figures of

BMI for 18-year-old girls from WHO[8]. The dotted lines in Figure 1(b) mark the reference cut-offs of thinness, underweight, median and overweight. We see that most ED-ed users have CBMI values lower than normal and their GBMI values are around the clinically underweight cut-off. These evidences demonstrate the effectiveness of our method in targeting ED-ed populations on Twitter.

6.2 Comparisons of User Features

We now present some descriptive analyses on the differences of ED-ed and non-ED-ed users, based on the measures we used in user characterization. Table 5 lists the mean and standard deviation values of some representative measures. We use the Kolmogorov-Smirnov (KS) test [31] to evaluate the statistical significance of differences between two sets of users, and use the Bonferroni correction to counteract the problem of multiple comparisons [23]. We see that most measures can distinguish well between ED-ed and non-ED-ed users. Comparing the KS statistics of different sets of users, the differences of ED-ed and non-ED-ed users are generally larger than those of Random and Younger users. This indicates that the sampled ED-ed users are significantly different from the general population on Twitter.

For social status, ED-ed users show the least social engagement, indicating that they have smaller #followees, #tweets and #followers than non-ED-ed users (see Eq. 1). However, ED-ed users do not show the least activity in the three sets of users. Basically, activity measures are the ratios of #followees, #tweets or #followers to active-period lengths (see Eq. 2). We thus conclude that ED-ed users are generally active on Twitter over a relatively shorter time period. For behavioral patterns, we find that ED-ed users prefer to post original tweets (i.e., %tweet) rather than re-tweeting others' tweets (i.e., %re-tweet and %quote). Besides, ED-ed users have less interactions with other users (e.g., %mention and %reply) and follow fewer public topics (e.g., %hashtag and %URL); their interactions with the external world are less diverse than those of non-ED-ed users. These results confirm that ED-ed users often suffer from social anxiety and they are shy to interact with others [26]. For psychometric properties, we see that ED-ed users use more of the 1st person singular (i.e., "I") and less of the 1st person plural (i.e., "we"), reflecting ED-ed users' loneliness, self-focused attention and psychologically distancing from others [18]. Also, ED-ed users express less positive emotion but more negative emotion (e.g., anger and sadness) in their posts, which may reflect their tendencies for depression, mental instability and irritability. Most of these indications are the common symptoms of ED [26]. Finally, we see that ED-ed users are more concerned about body image and ingestion, which is another important signal of ED [1, 4]. In contrast, non-ED-ed users care more about work and leisure than ED-ed users.

6.3 Classification Performance

Next, we assess the ability of using the above features to classify ED-ed and non-ED-ed users. Table 6 lists the mean and standard deviation values of four metrics for classification evaluations. We see that ED-ed users are clearly distinguishable from non-ED-ed users. Notably, the classification accuracy is above 97%, significantly higher than those in prior studies (e.g., 72% in predicting depression

---

[8]http://www.who.int/growthref/who2007_bmi_for_age/en/

Table 5: Statistics of measures characterizing differences of ED, Random and Younger users. Results of Kolmogorov-Smirnov tests comparing each pair of user sets (significance levels with Bonferroni correction: * $p < 0.01/m$; ** $p < 0.001/m$; *** $p < 0.0001/m$ where $m = 97$). Maximum values of each row are shown in bold.

Category	Measure	ED (c_1)	Random (c_2)	Younger (c_3)	$ks(c_1, c_2)$	$ks(c_1, c_3)$	$ks(c_2, c_3)$
Characteristics of Social Status							
Engagement	#Followees	4.86(σ=1.29)	**6.56**(σ=2.03)	6.17(σ=1.36)	**0.42*****	0.41***	0.14***
	#Tweets	5.20(σ=1.85)	**7.95**(σ=2.43)	6.72(σ=2.44)	**0.53*****	0.37***	0.24***
	#Followers	4.53(σ=1.52)	**7.54**(σ=2.82)	5.76(σ=1.83)	**0.53*****	0.33***	0.28***
Activity	#Followees/day	**1.11**(σ=1.02)	**1.11**(σ=1.29)	0.76(σ=1.02)	0.12***	**0.25*****	0.15***
	#Tweets/day	1.22(σ=0.84)	**1.92**(σ=1.36)	1.05(σ=0.93)	0.26***	0.13***	**0.31*****
	#Followers/day	0.91(σ=0.86)	**1.80**(σ=1.92)	0.60(σ=0.81)	0.23***	0.25***	**0.32*****
Characteristics of Behavioral Patterns							
Tweeting Preference	%Tweet	**0.74**(σ=0.21)	0.66(σ=0.29)	0.69(σ=0.27)	**0.15*****	0.09***	0.06***
	%Re-tweet	0.26(σ=0.21)	**0.34**(σ=0.29)	0.31(σ=0.27)	**0.15*****	0.09***	0.06***
	%Quote	0.00(σ=0.01)	**0.03**(σ=0.07)	0.01(σ=0.04)	**0.54*****	0.47***	0.10***
	%Mention	0.27(σ=0.19)	0.42(σ=0.28)	**0.44**(σ=0.26)	0.29***	**0.33*****	0.07***
	%Reply	0.08(σ=0.09)	0.13(σ=0.16)	**0.18**(σ=0.16)	0.17***	**0.35*****	0.19***
	%Hashtag	0.14(σ=0.15)	**0.23**(σ=0.27)	**0.23**(σ=0.21)	0.19***	**0.24*****	0.11***
	%URL	0.03(σ=0.09)	**0.26**(σ=0.28)	0.25(σ=0.26)	**0.64*****	0.63***	0.02***
Interaction Diversity	ΔRe-tweet	2.90(σ=1.47)	**4.39**(σ=1.58)	3.91(σ=1.51)	**0.41*****	0.31***	0.17***
	ΔMention	2.06(σ=1.24)	**3.37**(σ=1.46)	3.35(σ=1.44)	**0.41*****	**0.41*****	0.01
	ΔReply	1.94(σ=1.27)	**3.33**(σ=1.53)	3.14(σ=1.41)	**0.40*****	0.37***	0.06***
	ΔHashtag	2.65(σ=1.31)	3.85(σ=1.51)	**4.16**(σ=1.64)	0.38***	**0.46*****	0.12***
Characteristics of Psychometric Properties							
Linguistic Styles	1st pers singular	**0.10**(σ=0.03)	0.04(σ=0.03)	0.05(σ=0.03)	**0.76*****	0.74***	0.21***
	1st pers plural	0.00(σ=0.00)	**0.01**(σ=0.01)	**0.01**(σ=0.01)	0.61***	**0.62*****	0.08***
Affective Processes	Positive emotion	0.04(σ=0.01)	0.05(σ=0.02)	**0.06**(σ=0.02)	0.35***	**0.49*****	0.14***
	Negative emotion	**0.04**(σ=0.02)	0.02(σ=0.01)	0.02(σ=0.01)	0.54***	**0.64*****	0.11***
	Anger	**0.02**(σ=0.01)	0.01(σ=0.01)	0.01(σ=0.01)	0.38***	**0.53*****	0.15***
	Sadness	**0.01**(σ=0.01)	0.00(σ=0.00)	0.00(σ=0.00)	**0.64*****	0.62***	0.06***
Biological Processes	Biological processes	**0.07**(σ=0.02)	0.03(σ=0.02)	0.03(σ=0.01)	0.79***	**0.84*****	0.09***
	Body	**0.02**(σ=0.01)	0.01(σ=0.01)	0.01(σ=0.01)	0.62***	**0.68*****	0.07***
	Health	**0.02**(σ=0.01)	0.01(σ=0.01)	0.01(σ=0.01)	0.73***	**0.81*****	0.14***
	Ingestion	**0.04**(σ=0.02)	0.00(σ=0.01)	0.01(σ=0.01)	**0.86*****	0.85***	0.15***
Personal Concerns	Work	0.01(σ=0.01)	**0.02**(σ=0.01)	**0.02**(σ=0.01)	0.41***	**0.55*****	0.14***
	Leisure	0.01(σ=0.01)	**0.02**(σ=0.02)	**0.02**(σ=0.01)	0.39***	**0.62*****	0.26***

Table 6: Performance of predicting the classes of ED, Random (RD) and Younger (YG) users.

Measure	ED-RD	ED-YG	RD-YG
Accuracy	.972(σ=.036)	.982(σ=.011)	.793(σ=.029)
Precision	.982(σ=.017)	.986(σ=.007)	.797(σ=.028)
Recall	.972(σ=.036)	.982(σ=.011)	.793(σ=.029)
F1	.975(σ=.031)	.983(σ=.010)	.791(σ=.029)

[20]). There are two main reasons for such enhanced accuracy. First, self-reported diagnoses in users' profiles are more useful to accurately target positive users than identification on the basis of posts in prior studies [14]. Second, our sampling method enabled us to construct significantly more training samples than prior methods (e.g., 171 positive users in [20]). These results further confirm the effectiveness of our method in sampling ED-ed users on Twitter. Due to different sampling methods in use, we see that Random and Younger users can also be classified well, but with an accuracy lower than those in classifying ED-ed and non-ED-ed users. Another finding is that the differences of ED-ed and Younger users are larger than those of ED-ed and Random users, which is against our expectation. We conjecture the reason is because we gather Younger users by starting from popular artists. This may cause sampling biases and hence collect a set of users from some specific communities that are more different from ED-ed users than Random users.

We further examine the importance of each type of feature in predicting ED-ed users. We train one classifier each using: 1. social status; 2. behavioral patterns; 3. psychometric properties; 4. all features. Figure 2 shows the Receiver Operating Characteristic (ROC) curves generated by classifiers with different types of features. Comparing different classifications, we see that the differences between ED-ed and non-ED-ed users are consistently larger than those between Random and Younger users measured by each type of features. Comparing different types of features, although the best performance in each classification is achieved by using all features, psychometric properties alone are the best to distinguish different classes of users, particularly, achieving almost the same performance as using all features in classifying ED-ed and non-ED-ed users. This illustrates that: 1. the words people used in tweets are effective to reflect their mental health states; 2. we have sampled ED-ed users that are easily distinguishable from non-ED-ed users, so that using fewer features seems good enough to classify them well.

6.4 Characteristics of Networks

We now discuss the characteristics of different networks in ED-ed users. Table 7 lists the statistics of follow, re-tweet, reply and mention networks among ED-ed users. Each network is constructed by ED-ed users who have at least one corresponding link to other ED-ed users in our dataset, e.g., at least one who-follows-whom link in the follow network. All loop edges are ignored. Note that, due to the settings of Twitter API, all re-tweeters of a tweet in each cascade are
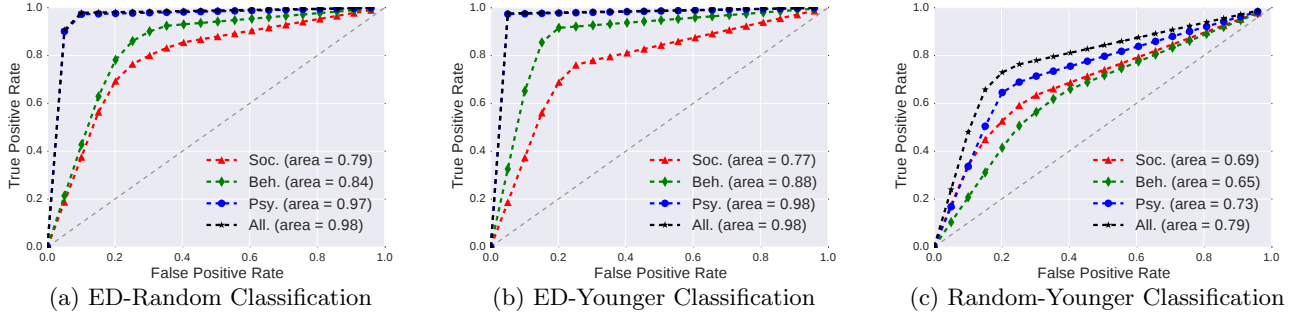
Figure 2: ROC curves of classifications with different types of features. The larger area under the curve (AUC) indicates the better performance. Gray dotted lines denote chance performance.

directly linked to the tweet's author in the re-tweet network. For example, if Bob re-tweets Andy and then Cole re-tweets Bob, both Bob and Cole are linked to Andy, even though Cole has not re-tweeted Andy directly. As most re-tweeting cascades are fairly shallow [5], all re-tweeters of a tweet can be regarded as direct re-tweeters approximately [44].

Table 7: Statistics of networks between ED-ed users.

Measure	Follow	Re-tweet	Reply	Mention
#Nodes	3,143	2,128	1,403	941
#Edges	52,982	11,338	4,344	1,408
Density	0.005	0.003	0.002	0.002
Avg. Path	3.156	4.398	5.030	5.258
#Components	2	10	22	37
%Giant Comp.	99.9	99.1	96.9	91.9
Clustering Coef.	0.122	0.089	0.052	0.029
Reciprocity	0.556	0.098	0.570	0.091
Degree Ass.	-0.105	0.002	-0.017	-0.199

Since Twitter follow networks have been intensively studied in the previous work [28, 34, 8], we first compare our feature statistics on the follow network with those in the literature. The average shortest path lengths in the ED follow network is 3.156 which is smaller than 4.05 and 4.8, the reference average shortest path lengths in the Twitter follow network reported in [34] and [8] respectively. This indicates that ED users connect one another more tightly than the general population on Twitter. As reported in [28], the reciprocity in the Twitter follow network is low at 0.221. However, the reciprocity in the ED follow network is 0.556, significantly higher than the reference reciprocity. This illustrates that ED-ed users have a relatively high density of social ties and have formed a tightly linked community. From the values of degree assortativity, we see that the follow network is disassortative by degree, i.e., users who have many followees are unlikely to be followed by others who have many followers, which aligns with the results in prior studies [34, 8]. Moreover, the degree assortativity in ED users (i.e., -0.105) is smaller than that in the general population (e.g., -0.0089 in [8]). This means that the more people an ED-ed user follows, the less popular the user tends to be.

We then explore the features of other ED networks. From the reciprocity of the reply network, we find that frequent mutual communications occur in tightly knit groups. These findings indicate that ED-ed users tend to engage in socializing and communicating with other ED-ed users on Twit-

ter. In conjunction with our previous findings that ED-ed users like to express negative emotions and discuss about body image and ingestion, we conjecture that ED-ed users may use Twitter to seek social support from other ED-ed peers and exchange ED-specific information [26]. Similar to the follow network, the reply and mention networks are also disassortative by degree. However, we find that the re-tweet network is assortative, i.e., users who have been re-tweeted a lot tend to re-tweet others who often re-tweet, which is in line with the results in [8]. That is, popular re-tweeters often seek information from other active re-tweeters. This sounds reasonable, as we can easily understand why information could propagate through Twitter by re-tweeting based on this [43]. Note that the statistics discussed above are potentially biased due to the bias of the data we collected.

6.5 Patterns of Homophily

Next, we present the results of homophily analysis in these networks. According to the significance test results of assortativity by each of 97 features, we list the percentages of features by which networks are assortatively mixed at different significance levels in Table 8. We see that various networks of ED-ed users are significantly assortative by most features, especially in the follow, re-tweet and reply networks. For example, for 85.6% of features, users with high feature values significantly tend to be connected to others with high feature values in the follow network, at significance level of 0.05. This indicates the presence of homophily in ED-ed communities. This, in turn, illustrates the feasibility of our snowball sampling ED-ed users through their follow networks.

Table 8: Percentages of assortatively mixed features at different significance levels.

Sig. Level	Follow	Re-tweet	Reply	Mention
$p < 0.05$	85.6%	83.5%	75.3%	46.4%
$p < 0.01$	79.4%	78.4%	58.8%	30.9%
$p < 0.001$	68.0%	69.1%	47.4%	17.5%

For a more detailed discussion, we rank these features by their values of z-score: $z = (r - \mu)/\sigma$, where r is the assortativity coefficient of networks by a feature, and μ and σ are the mean and standard deviation of the randomly simulated assortativity coefficients by the feature respectively. Table 9 shows the statistics of features ranked in the top 5 for each network. An interesting finding is that the feature "parenth", which denotes the percentage of using parentheses (e.g., '(',

Table 9: Examples of assortative mixing by features, ranked by the absolute values of z-score. Statistical significance tests are based on two-tailed hypothesis tests (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

Network	Feature	r	μ	σ	z	p
User Characteristics						
Follow	Parenth	.13	.00	.005	24.87	.0***
	%Quote	.12	.00	.005	24.63	.0***
	Death	.11	.00	.005	20.71	.0***
	%Tweet	.10	.00	.005	19.88	.0***
	%Re-tweet	.10	.00	.005	19.85	.0***
Re-tweet	ΔHashtag	.20	.00	.009	20.81	.0***
	%Quote	.16	.00	.009	18.29	.0***
	Parenth	.17	.00	.010	17.50	.0***
	Sad	.17	.00	.010	17.29	.0***
	SemiC	.13	.00	.009	14.25	.0***
Reply	%Quote	.21	.00	.016	13.42	.0***
	Parenth	.24	.00	.019	13.10	.0***
	%Tweet	.18	.00	.019	9.61	.0***
	%Re-tweet	.18	.00	.019	9.52	.0***
	Death	.15	.00	.019	8.02	.0***
Mention	Parenth	.25	.00	.028	8.84	.0***
	%Re-tweet	.22	.00	.028	7.94	.0***
	%Tweet	.22	.00	.028	7.89	.0***
	%Quote	.16	.00	.023	6.67	.002**
	%Reply	.16	.00	.028	5.66	.0***
Bio-indicators						
Follow	HW	.05	.00	.013	4.00	.001***
	LW	.09	-.01	.038	2.66	.021*
Reply	CW	.06	.00	.029	2.21	.038*

')'), is ranked very highly across different networks. To investigate this, we go through the posts of some users. We find that most parentheses are used to represent emoticons, such as ":))", ":((". That is, users who like to use emoticons tend to connect with others who also like to use emoticons. This means that ED-ed users have similar habits in using language. Other significantly assortative features include tweeting preferences (e.g., %tweet and %quote), diversities of using hashtags, concerns of death, and emotion (e.g., sad).

Moreover, we employ a similar method to investigate the homophily of ED-ed users in terms of their bio-information indicators (in Table 4). The results are listed in the bottom of Table 9. We find that ED-ed users tend to follow others who have similar HW or LW, and tend to reply to others who have similar CW. In other words, ED-ed users often seek acquaintances with others who have similar experiences on weight management, while they communicate with others who are in a similar situation at present. Thus, we suppose that ED-ed users follow others to seek a sense of community identity and peer support, and reply to others perhaps to discuss weight loss or other contingent information.

7. CONCLUSION AND FUTURE WORK

In this paper, we study to detect and characterize ED-ed communities on SM. We first present a snowball sampling method to automatically sift ED-ed individuals and their community structures from Twitter data. We then compare ED-ed and two sets of non-ED-ed users in social status, behavioral patterns and psychometric properties, and find that ED-ed users show young ages, prevailing urges to lose weight even if being clinically underweight, high social anxiety, intensive self-focused attention, deep negative emotion, increased mental instability, and excessive concerns of body image and ingestion on Twitter. We further build classifiers to classify ED-ed and non-ED-ed users, and show that Twitter data can help estimating the occurrence of ED. Finally, we leverage the social networking data among ED-ed users, and present the first empirical homophily analysis of ED-ed communities on SM. We find that: 1. ED-ed users have significant assortative mixing patterns in tweeting preferences, language use, concerns of death and emotions etc.; 2. ED-ed users tend to follow and reply to other ED-ed users having similar body weight. Our findings indicate that ED-ed individuals primarily use SM for a sense of community identity and mutual social support online. Moreover, the presence of homophily in ED-ed communities and the accuracy of more than 97% in predicting ED-ed users show the feasibility of develop computational methods to detect larger ED-ed communities on Twitter, beyond those that have self-identified.

In the future, we intend to gain a deeper understanding of the features of ED-ed communities on SM, e.g., automatically-learned topics (e.g., using topic modeling) in their posts and individuals' opinions towards ED. We also plan to examine whether we can use our method to study other mental illnesses on other SM sites; whether users who self-identify with a mental illness differ from others who have not self-identified. Another direction is to study whether node and graph properties (e.g., centrality) of social networks among mentally ill individuals relate to the severity of illness scoring; (how) can people's peers on SM influence their mental health states over social networks. Pursing the answers to these questions will be valuable in the design of interventions for large communities affected by a mental illness.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] D. S. Abebe, L. Lien, and T. von Soest. The development of bulimic symptoms from adolescence to young adulthood in females and males: A population-based longitudinal cohort study. *International Journal of Eating Disorders*, 45(6):737–745, 2012.

[2] J. Arcelus, A. J. Mitchell, J. Wales, and S. Nielsen. Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Archives of general psychiatry*, 68(7):724–731, 2011.

[3] A. Arseniev-Koehler, H. Lee, T. McCormick, and M. A. Moreno. # proana: Pro-eating disorder socialization on twitter. *Journal of Adolescent Health*, 58(6):659–664, 2016.

[4] D.-. A. P. Association et al. Diagnostic and statistical manual of mental disorders. *Arlington: American Psychiatric Publishing*, 2013.

[5] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts. Everyone's an influencer: quantifying influence on twitter. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 65–74. ACM, 2011.

[6] N. N. Bazarova and Y. H. Choi. Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of Communication*, 64(4):635–657, 2014.

[7] Beat. The costs of eating disorders: Social, health and economic impacts. From https://www.b-eat.co.uk/assets/000/000/302/The_costs_of_eating_disorders_Final_original.pdf?1424694814, 2015-02.

[8] D. R. Bild, Y. Liu, R. P. Dick, Z. M. Mao, and D. S. Wallach. Aggregate characterization of user behavior in twitter and analysis of the retweet graph. *ACM*

Transactions on Internet Technology (TOIT), 15(1):4, 2015.

[9] S. Chancellor, Z. Lin, and M. De Choudhury. "this post will just get taken down": Characterizing removed pro-eating disorder social media content. In *Proceedings of the 2016 CHI Conference*. ACM, 2016.

[10] S. Chancellor, Z. Lin, E. L. Goodman, S. Zerwas, and M. De Choudhury. Quantifying and predicting mental illness severity in online pro-eating disorder communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, pages 1171–1184. ACM, 2016.

[11] S. Chancellor, T. Mitra, and M. De Choudhury. Recovery amid pro-anorexia: Analysis of recovery in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2111–2123. ACM, 2016.

[12] S. Chancellor, J. A. Pater, T. Clear, E. Gilbert, and M. De Choudhury. # thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, pages 1201–1213. ACM, 2016.

[13] M. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonçalves, F. Menczer, and A. Flammini. Political polarization on twitter. *ICWSM*, 133:89–96, 2011.

[14] G. Coppersmith, C. Harman, and M. Dredze. Measuring post traumatic stress disorder in twitter. In *ICWSM*, 2014.

[15] A. Culotta. Estimating county health statistics with twitter. In *Proceedings of the 32nd CHI conference*, pages 1335–1344. ACM, 2014.

[16] M. De Choudhury. Anorexia on tumblr: A characterization study. In *Proceedings of the 5th International Conference on Digital Health 2015*, pages 43–50. ACM, 2015.

[17] M. De Choudhury, S. Counts, and E. Horvitz. Major life changes and behavioral markers in social media: case of childbirth. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1431–1442. ACM, 2013.

[18] M. De Choudhury, S. Counts, and E. Horvitz. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 47–56. ACM, 2013.

[19] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff. Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM conference on Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, pages 626–638. ACM, 2014.

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

[21] N. Eagle, M. Macy, and R. Claxton. Network diversity and economic development. *Science*, 328(5981):1029–1031, 2010.

[22] M. J. Esplen, P. Garfinkel, and R. Gallop. Relationship between self-soothing, aloneness, and evocative memory in bulimia nervosa. *International Journal of Eating Disorders*, 27(1):96–100, 2000.

[23] A. V. Frane. Are per-family type i error rates relevant in social and behavioral science? *Journal of Modern Applied Statistical Methods*, 14(1):5, 2015.

[24] A. S. Guarda. Treatment of anorexia nervosa: insights and obstacles. *Physiology & Behavior*, 94(1):113–120, 2008.

[25] G. C. M. D. C. Harman. Quantifying mental health signals in twitter. *ACL 2014*, page 51, 2014.

[26] A. S. Juarascio, A. Shoaib, and C. A. Timko. Pro-eating disorder communities on social networking sites: a content analysis. *Eating disorders*, 18(5):393–407, 2010.

[27] I. Kawachi and L. F. Berkman. Social ties and mental health. *Journal of Urban health*, 78(3):458–467, 2001.

[28] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages

[29] A. Lenhart, M. Duggan, A. Perrin, R. Stepler, L. Rainie, and K. Parker. Teens, social media & technology overview 2015, pew research center, 2015.

[30] K. Lerman, X. Yan, and X.-Z. Wu. The majority illusion in social networks. *arXiv preprint arXiv:1506.03022*, 2015.

[31] H. W. Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318):399–402, 1967.

[32] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.

[33] P. Metaxas, E. Mustafaraj, K. Wong, L. Zeng, M. OâĂŹKeefe, and S. Finn. What do retweets indicate? results from user survey and meta-review of research. In *Ninth International AAAI Conference on Web and Social Media*. Citeseer, 2015.

[34] S. A. Myers, A. Sharma, P. Gupta, and J. Lin. Information network or social network?: the structure of the twitter follow graph. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 493–498. ACM, 2014.

[35] M. E. Newman. Mixing patterns in networks. *Physical Review E*, 67(2):026126, 2003.

[36] N. I. of Mental Health. Eating disorders. Retrieved 2016-06-30 from http://www.nimh.nih.gov/health/topics/eating-disorders/index.shtml.

[37] S. J. Paxton, H. K. Schutz, E. H. Wertheim, and S. L. Muir. Friendship clique and peer influences on body image concerns, dietary restraint, extreme weight-loss behaviors, and binge eating in adolescent girls. *Journal of abnormal psychology*, 108(2):255, 1999.

[38] D. Rao, D. Yarowsky, A. Shreevats, and M. Gupta. Classifying latent user attributes in twitter. In *Proceedings of the 2nd international workshop on Search and mining user-generated contents*, pages 37–44. ACM, 2010.

[39] B. Raskutti and A. Kowalczyk. Extreme re-balancing for svms: a case study. *ACM Sigkdd Explorations Newsletter*, 6(1):60–69, 2004.

[40] S. Swan and B. Andrews. The relationship between shame, eating disorders and disclosure in treatment. *British journal of clinical psychology*, 42(4):367–378, 2003.

[41] S. Syed-Abdul, L. Fernandez-Luque, W.-S. Jian, Y.-C. Li, S. Crain, M.-H. Hsu, Y.-C. Wang, D. Khandregzen, E. Chuluunbaatar, P. A. Nguyen, et al. Misleading health-related information promoted through video-based social media: anorexia on youtube. *Journal of medical Internet research*, 15(2):e30, 2013.

[42] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54, 2010.

[43] I. Taxidou and P. M. Fischer. Online analysis of information diffusion in twitter. In *Proceedings of the 23rd International Conference on World Wide Web*. ACM, 2014.

[44] L. Weng and F. Menczer. Topicality and impact in social media: Diverse messages, focused messengers. *PloS one*, 10(2):e0118410, 2015.

[45] M. Wolf, F. Theis, and H. Kordy. Language use in eating disorder blogs: Psychological implications of social online activity. *Journal of Language and Social Psychology*, page 0261927X12474278, 2013.

[46] I. Wood. Using topic models to measure social psychological characteristics in online social media. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pages 308–313. Springer, 2015.

[47] E. Yom-Tov, L. Fernandez-Luque, I. Weber, and S. P. Crain. Pro-anorexia and pro-recovery photo sharing: a tale of two warring tribes. *Journal of medical Internet research*, 14(6):e151, 2012.