Cultivating the Community: Inferring Influence Within Eating Disorder Networks on Twitter

FAYIKA FARHAT NOVA, Marquette University, USA

AMANDA COUPE, Parkview Research Center, USA

ELIZABETH D. MYNATT, Georgia Institute of Technology, USA

SHION GUHA, University of Toronto, Canada

JESSICA A. PATER, Parkview Research Center, USA

A growing body of HCI research has sought to understand how online networks are utilized in the adoption and maintenance of disordered activities and behaviors associated with mental illness, including eating habits. However, individual-level influences over discrete online eating disorder (ED) communities are not yet well understood. This study reports results from a comprehensive network and content analysis (combining computational topic modeling and qualitative thematic analysis) of over 32,000 public tweets collected using popular ED-related hashtags during May 2020. Our findings indicate that this ED network in Twitter consists of multiple smaller ED communities where a majority of the nodes are exposed to unhealthy ED contents through retweeting certain influential central nodes. The emergence of novel linguistic indicators and trends (e.g., “#meanspo”) also demonstrates the evolving nature of the ED network. This paper contextualizes ED influence in online communities through node-level participation and engagement, as well as relates emerging ED contents with established online behaviors, such as self-harassment.

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

Additional Key Words and Phrases: Eating Disorder; Influence; Online Communities; Meanspo; Self-harassment; Network Analysis

ACM Reference Format:

Fayika Farhat Nova, Amanda Coupe, Elizabeth D. Mynatt, Shion Guha, and Jessica A. Pater. 2022. Cultivating the Community: Inferring Influence Within Eating Disorder Networks on Twitter. *Proc. ACM Hum.-Comput. Interact.* 6, GROUP, Article 7 (January 2022), 33 pages. <https://doi.org/10.1145/3492826>

1 INTRODUCTION

Social media platforms are one of the most robust sources for peer-generated health information as they provide a platform for individuals to converse with one another about different health topics [146]. Various communities within platforms, such as Twitter, allow individuals with chronic health conditions to share data, provide and receive help, raise awareness, challenge stigma, document their experiences, and advocate for themselves [16, 22]. While participation in these communities can

Authors' addresses: Fayika Farhat Nova, fayikafarhat.nova@marquette.edu, Marquette University, Milwaukee, WI, USA; Amanda Coupe, amanda.coupe@parkview.com, Parkview Research Center, Fort Wayne, IN, USA; Elizabeth D. Mynatt, mynatt@cc.gatech.edu, Georgia Institute of Technology, Atlanta, GA, USA; Shion Guha, shion.guha@utoronto.ca, University of Toronto, Toronto, ON, Canada; Jessica A. Pater, jessica.pater@parkview.com, Parkview Research Center, Fort Wayne, IN, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-0142/2022/1-ART7

<https://doi.org/10.1145/3492826>

positively impact users, there is potential for harm as well. For example, there are many online pro-Eating Disorder (ED) communities that actively promote ED activities and behaviors as a legitimate lifestyle choice rather than an expression of a serious and pervasive chronic mental illness [92, 102]. EDs are often “hidden” due to cultural/societal norms valuing thinness; this can delay diagnosis and treatment, at times resulting in fatal consequences [53]. Online pro-ED communities increase these risky consequences and negatively affect people with and without EDs, via promoting unrealistic standards of thinness, encouraging disordered eating behaviors, and sharing harmful tips on how to develop and maintain an unhealthy weight (known as “thinspiration” or “thinspo”) [20, 136].

Twitter is a popular social media platform within ED communities as it provides a ubiquitous and open platform for pro-ED activity and content [11, 143]. HCI research using Twitter public data within this domain includes identifying and understanding eating disordered behaviors and activities online [99, 107, 143], predicting risks of ED activities [28], and discovering the in situ evolution of social norms and practices within online platforms as they relate to ED behaviors [30, 59]. Studies examining the spread of information through online communities have shown that the social structure of the network influences how far and fast new information spreads [25, 145]. For online pro-ED communities, influential nodes play an important role as they are often key to the vitality of the community and provide a representation of accepted behaviors and activities [119]. Unfortunately, this often means the glamorization and normalization of disease-related content that can create and/or exacerbate underlying health issues. However, research focused on node-level influence and nuanced community structure within platform-specific ED community networks online is limited.

Therefore, this study aims to address the gap in the existing literature and add insights on online pro-ED networks through analyzing the community structures and associated activities/behaviors that are particularly influenced by the central nodes. For this study, we report the findings of both qualitative and quantitative network analyses of various ED communities on Twitter. Our dataset was constructed with over approximately 32,000 data points collected through Twitter’s official premium API between May 1st-31st, 2020, using specific ED-related hashtags (#thinspo, #proana, #edtw). Using quantitative measures such as centrality matrix, rank correlation, and community detection algorithms (Louvain [134] and Girvan-Newman [47]) on this dataset, we identified central nodes influencing the ED communities and how the community is structured around them. Using combined computational topic modeling (i.e., LDA [114] and NMF [24] topic modeling algorithms) and qualitative thematic analysis, we also analyzed the text entities of the dataset (user tweets) to identify emerging and popular contents and topics within the ED network. Our findings show:

- The ED community on Twitter is a homophilic network where many interactions between nodes are done through retweeting certain influential nodes. This pattern indicates an information dissemination characteristic of the community, in which majority nodes are exposed to unhealthy ED contents through retweeting specific central nodes.
- Our community detection algorithms confirm the existence of defined communities and sub-communities within the Twitter ED network that are primarily led by these central nodes. However, depending on the structure of the network (whole vs. sub-graph), the shape of this leadership can be sparse.
- Along with existing popular contents, we observe emergence of new contents, such as meanspo-related tweets, which highlight the importance of receiving harsh or negative “encouragement” online for some individuals with EDs. Our analysis on the central nodes confirms that these popular contents are also common and significant within these influential nodes.

- We also identified emerging ED-related linguistic indicators based on topic modeling, including important primary and secondary themes such as community building, community curating, desired ED lifestyle, and emergence of novel community platforms. Such results indicate that community sense within ED individuals in Twitter is very strong, and it often became a central topic of discussion and participation.

This research makes three key contributions to the HCI literature: 1) We contextualize the presence of ED influence in online communities through node-level participation and engagement; 2) We provide a granular understanding of prominent topics and contents within ED community that extend our current knowledge on online ED networks and provide better contexts on their internal activities; and 3) We contextualize emerging ED-related linguistic indicators (e.g., meanspo) with known ED activities that are centered around self-harassment, and censorship and spark discussions on potential ways to translate this knowledge into clinical practice.

2 LITERATURE

2.1 Social Network Analysis

Network analysis of online platforms maps and measures the relationships and flows between individuals, groups, organizations, URLs, and other information and knowledge entities linked to them [85]. Twitter plays a key role in information dissemination online [79], and the structure of a Twitter network has a substantial impact on information dissemination among users [73]. Community detection is the first step in understanding a network's underlying semantic and structural information [5] and by using quantitative measures of social network analysis (SNA), we can identify significant influential individuals (centrality measures) who are critical for the distribution of potentially useful and/or harmful contents [90] and examine their interactions with other community members [132, 151].

Influencer detection is an important aspect of SNA, as they can create a chain-reaction of influence that reaches a very large scale of users [72] and are one of the main characters of information diffusion in online social networks [94]. It is not unusual for a small group of influential and active community members to provide the majority of online content [7] and influence health-related behavior among peers [83]. For example, a study on children's food habits showed influencers on YouTube promote unhealthy food and drinks to their followers [8]. Relatedly, Chin et al. [36] worked on YouTube's network structure to understand how information on vaccination disseminates [151]. [37], using logistic regressions with centrality measures, evaluated the social network structure of QuitNet Smoking Cessation and identified influential subgroups whose existence and characteristics might inform the design of cessation interventions. In the broader field of health information distribution, Hanson et al. [63] highlight the importance of neighbors (network effect) on user engagement in Twitter related to drug abuse [124]. These research on online social networks indicate individual adoption of certain behaviors/activities will more likely to happen when participants received social reinforcement from multiple prominent neighbors in the social network.

2.2 ED Related Studies

Eating disorders have a significant impact on both mental and physical health and affect around 9% of individuals at some point in their lifetimes, resulting in a death every 52 minutes [53], which is the highest mortality rate of any mental disorder [10]. People with ED tend to have a high need for social approval and a strong fear of rejection; thus, EDs are associated with body dysmorphia/impression management concerns [117], a core concept in HCI [62]. Evidence from multiple studies examining pro-ED contents shared online have also reported that exposure to such information causes higher body dissatisfaction, increased eating disturbance, lower self-esteem,

and a poorer body image [14, 64, 131]. Online pro-ED communities are often hidden in plain sight, with their actions walled off from the mainstream of users' activities, yet are easily accessible by searching for related tags/keywords [30].

Existing research has tried to identify and analyze these tags/keywords [107], along with linguistic variations [30, 80], censorship around ED participation [27], and the ethical implications of pro-ED materials in online groups [125] using both qualitative (codebook) and quantitative (topic modeling) content analysis methods. Many of these studies are Twitter-specific, as it provides a sense of community identity and mutual social support [143] by allowing users to publicly exchange harmful pro-ED contents through unique "follow," "retweet," "reply," and "mention" features [11, 143]. Such interactions online encourage the sharing of ED-related knowledge, attitudes, and behaviors with the larger network, amplifying the destructive impact community members have on themselves [107]. Although there has been extensive work on identifying signs of ED from user-generated content on Twitter, few have investigated the community membership and structure of these heterogeneous networks in depth [30].

As an attempt to fill that gap, [143] focused on users' ED participation on Twitter, and based on interactions between individuals, presented the existence of homophily among ED communities on social media. Others have described the structure and evolution of communication in online ED communities using clustering techniques to identify ED-related issues discussed in online conversations and to depict interpersonal connections in the network [144]. Additional research includes comparing ED communities (#thinspiration vs. #fitspiration) to observe the information flow within users [138] and investigating senders and associated tweets within #proana communities [65]. While these findings are insightful, they largely focus on community connections without examining individual influence within ED networks and its impact on the creation, distribution, and adoption of harmful information within mass populations online.

2.3 Negative Impacts of Life Online

When discussing the scope of peer influence within online communities, one must consider the pervasiveness of social media platforms within daily life in the U.S. [12]. Extant literature has examined a wide range of both positive and negative effects of life online. For example, individuals are able to connect and maintain relationships with friends, family members, and colleagues [123, 140] and develop new ones [34, 66]. However, social media and other forms of computer-mediated communication are also associated with harmful interpersonal interactions (e.g., harassment [17, 69], cyberbullying [109, 128], doxxing [52, 130], trolling), non-consensual and/or age-inappropriate exposure to explicit sexual content or behaviors (e.g., sexual solicitation [148], sharing of nude images [39, 109]), development or exacerbation of addiction issues (e.g., internet addiction [147]), and the erosion of privacy related to personal content online [86, 129].

With the ubiquity of smartphones and constant connectedness in America [9], it is expected that many teens and young adults have engaged in and/or been exposed to risk behaviors online. For example, the majority of adolescents report either being cyberbullied or witnessing cyberbullying [31, 109]. Research suggests that adolescents experience negative emotional effects after single incidences of exposure to certain online risks and behaviors (i.e., cyberbullying, sexual solicitation, explicit content), but are able to recover quickly; however, this does not account for repeated or consistent longitudinal exposure, nor does it consider its potential developmental effects [87]. Relatedly, many adolescents lack an accurate or nuanced understanding of important concepts related to internet and social media usage, such as privacy and "real life" effects of online behavior [109]. As it is estimated that 90% of EDs are diagnosed before the age 20 [4], these age-related social media concerns may hold particular relevance for online ED communities.

2.4 Mental Health and HCI

Mental health and HCI is a growing body of research that has focused on various domains and the application of various computing techniques to uncover [19, 20], characterize [107, 110], and predict [28, 68, 113] mental health issues online, misuse of wearables [48, 54], and the development of platforms for ongoing support [76, 100, 139], to name a few. Research has gone beyond the specific disease states to understand and make inferences about technology and hospitalizations [56] as well as treatment [105, 137]. In the last decade, research has diversified from the more explored depression [42, 51] and anxiety [89, 127] to include non-suicidal self-injury [75], eating disorders [30, 107, 143], suicidality [96], bipolar [38], and schizophrenia [57, 142]. Additionally, studies have examined gender and culture as they relate to online activities [44, 81, 110]. Pater et al. conducted a mixed methods case study including a review of the social media activity of eating disorder patients during the 18 months prior to their first hospitalization [106]. They found that patients were avid consumers of online ED content, but not producers. This research raised questions about who the influencers within the ED networks are and what type of content are they sharing.

Thus, this study aims to add insights on the community structure of online Eating Disorder networks and the associated unhealthy and dangerous activities and behaviors that are particularly influenced by the key members. We chose Twitter for this analysis because it is one of the most well-known and well-used platforms within ED communities. Our research is grounded in these important questions:

RQ1: What does the community structure look like for ED networks on Twitter and who are influencing these online ED communities?

RQ2: What types of ED contents (hashtags, media, links) are popular and shared within these ED networks on Twitter?

RQ3: What emerging ED-related linguistic indicators can we identify from the ED communities on Twitter?

3 METHODS

In this section, we provide a detailed description of the Twitter dataset along with the data collection methods, preprocessing techniques, and qualitative analysis. We subsequently describe the community and topic detection algorithms used in this study to identify communities and prominent topics within our dataset. We also describe the codebook for our analyses of the hashtags, text, and media content contained within posts.

3.1 Study Data

We used Twitter's official API to collect over 32,000 public posts ("tweets") in the pro-ED space between May 1st and May 31st, 2020. As Twitter's free API only collects public tweets from the last 7 days, we used paid premium full archive API to gather a month's worth of targeted tweets [1]. The sample size and timeframe used in this study are consistent with similar studies by domain experts [30, 43]. Our data collection methodology followed an established approach [30] and occurred in three stages. First, we obtained posts counts using the primary tag #thinspo, which is known to be highly related to eating disorder content [30, 107]. The resulting sample of approximately 14,000 tweets provided a list of top tags connected to the #thinspo community within Twitter. We then collected additional tweets using the two most common tags from this initial list (#proana and #edtwit). Table 1 provides an overview of the final dataset created from the three ED-related tags. Finally, we created a candidate set of posts from these raw sets that we confirmed to be related to pro-ED behavior.

Number of Total Tweets	32530
Number of Unique Users	5658
Total Retweets	19165
Actual Tweets	13365
Total Tweets with Links	8262
Mean Tweets Per User	6.67
Variance of number of posts per day per user	12.42

Table 1. Statistics of Twitter data of study cohort

Prior to topic modeling, we pre-processed the tweets in a manner consistent with similar prior work [134]. We did not alter hashtags (starting with “#”) or usernames (starting with “@”), but transformed all remaining text to lowercase and eliminated stopwords, the term “RT” (which stands for “retweet”), and numerical values. The Porter stemmer [71] was then used. We excluded URLs (uniform resource locators) generated by URL shortening services (e.g., “http://bit.ly”) and included the domain of any full URLs identified from the list of expanded URLs. The final dataset included over 32,000 tweets from 5,658 users, including over 19,000 retweets (tagged as RT) and 1,141 unique tags. A few sample posts are shown in Table 2. The variables collected for each tweet included:

- **Timeline:** The timeline of the tweet posted (integer)
- **ID:** Unique ID of the user (integer)
- **Name:** Name of the user (string data type)
- **Screen name:** Unique screen name of the user (string data type)
- **Reply count:** How many replies each tweet received (integer)
- **Retweet count:** How many retweets each tweet received (integer)
- **Like count:** How many likes each tweet received (integer)
- **Total Status:** How many total tweets that user has shared at the time of posting (integer)
- **Friends Count:** How many users the user follows at the time of posting (integer)
- **Followers Count:** How many total followers the user has at the time of posting (integer)
- **Tweets:** The individual tweets that included text, hashtags, and media links (string)

I am sure being skinny will make me happiest and I won't stop until I reach there #thinspo
Don't stop trying until you become the #thinspo for others!
RT: hi..jst joined #edtw! let's be moots and rt each other
Okay..gotta up my game...no overeating at any cost!#edtw
If I take pics of my weight daily and upload in #edtw, would you guys join?
RT: smoking ruins my appetite...what's your secret to battle hunger? #proana #thinspo

Table 2. Example posts from users in the dataset

We used undirected graphs to create two networks (connected through only either mentions or retweets) using the 10,709 connections between the 4,786 users. Each node represents a user/vertex, and an edge between two users is established if they were connected through RT/mention. To identify key nodes within the networks, we used a centrality matrix [150]. In graph theory and network analysis, centrality of a vertex measures its relative importance within the graph. Members with the highest structural relevance are identified by high centrality ratings and are assumed to perform key roles in the network [150]. For community detection, Girvan-Newman (GN) [47] and Louvain [134] algorithms were used. We chose GN algorithm due to its use of edge betweenness

(a measure of the centrality and influence of an edge in a network) [91] and Louvain algorithm for its wide popularity across multiple domains, as well as its low time complexity and sequential access feature performance [73, 115]. We ran Latent Dirichlet Allocation (LDA) [114] and the Non-Negative Matrix Factorization (NMF) [24] algorithms, two widely popular algorithms used within similar domains, to infer topics from the post-processed tweets [95, 134]. We then mapped the topics back onto the whole set, such that each tweet was connected with a specific topic.

3.2 Community Detection

Community detection algorithms identify sets of nodes in a graph with a higher density of community connections within themselves (i.e., in one set) than between multiple sets [134].

The Girvan-Newman algorithm detects communities within large, complex networks by using the concept of "edge betweenness" to continuously remove the edges that are not relevant to a particular user [135]. It follows these steps [47]:

- 1) Calculate edge betweenness for every edge in the graph
- 2) Remove the edge with highest edge betweenness
- 3) Calculate edge betweenness for remaining edges
- 4) Repeat steps 2-4 until all edges are removed

To calculate edge betweenness, the algorithm finds the shortest paths in the graph, starting with one vertex and calculating edge weights for paths that pass through it, then repeating the process for each vertex and adding the weights for each edge [47]. Because of its simplicity and ease of implementation, this algorithm has been successfully applied to a variety of networks [135]. However, because of its computational complexity, which is $O(m^2n)$ for a network with n nodes and m edges, it can only be used in networks with a few thousand nodes [135]. We used python implementation of `edge_betweenness` for GN algorithm.

The Louvain is more efficient than other algorithms and detects communities in large networks by maximizing modularity [18, 101, 116]. Using an agglomerative approach, nodes are first assigned to a community of size one; they are then successively aggregated with the neighboring community that yields the highest gain in modularity, if it exists [134]. The communities detected then become nodes in a new network, with edge weights defined by the number of connections between them [134]. The algorithm therefore constructs a hierarchical representation of the network and proceeds until no more modularity gains can be identified. This results final clustering that is used to define the community structure [134]. To find when an optimal decomposition of a network into communities is reached, modularity (Q) of the network is calculated in every step of divisible algorithms, and the maximum value q gives us the best partition of the graph. We used python package `community` and `module best_partition` for Louvain algorithm.

3.3 Topic Modeling Procedure

To perform unsupervised topic modeling on our dataset, we used LDA and NMF models. LDA is widely utilized algorithm widely utilized in a range of fields for topic modeling text corpora, which may be utilized as a sort of text summary of a huge number of documents [114]. NMF differs from other dimensionality reduction methods by uncovering hidden low-dimensional patterns in high-dimensional data, providing a nonnegative, part-based representation [24]. Using algorithms that account for different types of interpretations allows for a more nuanced and comprehensive assessment of the dataset.

For both LDA and NMF, the topic modeling method is a multi-step technique [133]. Before applying the algorithms, we used standard text processing techniques to clean the data, including **tokenization** (representing each tweet with a sequence of tokens, i.e., words in the sense of the "bag-of-words" model), **normalization** (normalizing the sequence of tokens to a limited character set, i.e.,

Fig. 1. Flowchart representing the steps taken to detect topics using LDA and NMF

[a-z]), **URL removal**, **stopword removal** (deleting articles, conjunctions, prepositions, etc.), and **stemming** (reducing each word to its root form by a standard stemming algorithm). We filtered out tokens that were present in less than 15 documents (too rare to register similarity among documents) or more than 50% of the documents (too frequent to register difference among documents). We then kept only the first 100,000 most frequent tokens. We then completed vectorization of the data to document-term frequency matrix. In LDA, the model is trained over a corpus (i.e., the collection of words in the documents of our textual data). For NMF, the created document-term frequency is factorized as per the algorithm before printing of the obtained topics. The number of topics were decided based on repeated experimentation, as choosing too few topics causes vagueness, but too many results in overfitting [114]. After repeated trials of the same corpus with multiple topic sizes, we discovered that constructing models with ten topics resulted in best coherent values (ranging from 0 to 1) for our dataset. The final output for both LDA and NMF contained the top 20 key phrases and associated tweets for each relevant topic, as well as their contribution.

A typical complexity of LDA topic modeling is that the topics created are not necessarily clear or interpretable by humans [32]. While topics generally return a ranked list of the most common keywords, the ability to classify topics based on these keywords is limited because phrases used frequently across the corpus may rank high in numerous topic lists [114]. The result shows the relevance of each topic to each tweet in the dataset while also generating top keywords for each topic. According to our findings, NMF outperformed LDA in selecting themes from the dataset based on their coherence score (detailed further in results). A detailed and in-sequence picture of these processes are depicted in Fig. 1.

3.4 Codebook Analysis

Our dataset included public tweets related to the tags #thinspo, #edtwit, and #proana. The tweets included at least one of these components: hashtags, text, and attached media (i.e., image, video, gif). A total of 75,676 hashtags were attached to the 32,530 posts in our dataset, of which 1,141 were unique. We used an inductive approach to analyze the tags and identified common as well as emerging hashtag categories within our dataset. We also used an inductive approach to analyze the media content associated with 710 random posts within our dataset. Based on the link types, general themes were created and further refined into coding taxonomy. We also created media archetypes (the recurrence of characters or ideas sharing similar traits) [107] based on the types of

information the media links provided. Additionally, we conducted qualitative assessment of the topics generated from LDA and NMF to group and categorize similar themes.

3.5 Ethical Considerations

When conducting research with vulnerable online communities, we as researchers have attempted to keep in mind our own group memberships, identities, and potential lack of knowledge about these communities' experiences, challenges, and values. We note the ethical limitations of conducting this research: we used public datasets, which are not defined as human subjects research, and thus, we collected this data without user knowledge or consent. While this is an accepted practice when dealing with public data, future research may benefit from a collaborative approach with members of online communities of interest. Presently, this research may misrepresent the behaviors, challenges, or identities of the study population, and our methods do not allow us to clarify potential misrepresentations. Additionally, the social media activity in our analysis is unable to capture the many complexities and nuances of human behavior.

To ensure anonymity of the users in our dataset, we have replaced all usernames with pseudonyms. We have also used representative media and text examples and/or edited any direct quotes to ensure that the user-generated content reported in this paper cannot be searched and connected to the account that posted it.

While working in domains of a sensitive nature and conducting qualitative assessments, we believe it is important to disclose the positionality of the authors. The authors of this paper have domain expertise in psychology, network analysis, and online characterizations of mental health.

4 RESULTS

The aim of this paper was to explore the presence of ED influence in online communities through node-level participation and engagement, while also examining prominent topics and contents that emerge through quantitative and qualitative analysis.

4.1 Influence within ED Network

Of the approximately 32,000 tweets in our dataset, around 13,000 are original tweets (including mentions) and around 19,000 are retweets. To study characteristics of the central nodes, we translated the raw data into a particular network. More specifically, first, we formed a network of mentions where where if User A mentioned User B, we connected A to B using an edge [42]. Second, we formed a network of retweets where we connected each user to the other user(s) that they have retweeted [143]. As most retweeting cascades are fairly shallow [13], all retweeters of a tweet can be regarded as direct retweeters approximately [143]. The graph created from direct mentions is much smaller than the retweet graph. Whereas our retweet network (FG) is comprised of 4,747 unique nodes (denoted as U_r) with total 10,686 edges between them (denoted as E_r), the direct mention graph has only 39 unique nodes (denoted as U_m) with total 23 edges (denoted as E_m) between them. Even though the dataset is collected from a specific timeline, this very high deviation between these network structures in terms of nodes and edges is noteworthy, primarily indicating that users within our dataset are more prone to retweets than creating/sharing actual contents that includes mentions.

4.1.1 User Influence. To identify influence within the ED network, we decided to explore the existence of prominent nodes. To detect influential users in the network dynamics, we considered retweets as a proxy to represent an endorsement to the tweet content shared by the user. For analysis purposes, along with our graph with 4,747 unique nodes (for graph FG), we also looked at a subset of this network that only include nodes having 20 or more connections. Fig. 2 illustrates

the degree distributions of both graphs where only a few nodes have higher degree (2a) and the tails of the distributions follow a power law. The resulting network is represented as $SG = \langle V_{sg}, E_{sg} \rangle$, where V_{sg} are the set of nodes represented from U_r with E_{sg} edges between them. After filtering, the final graph SG consisted of a much smaller network with 179 nodes and 1061 edges.

Fig. 2. Degree Distributions of nodes from FG and SG

We conducted an analysis of centrality matrix to identify the most important nodes within the network. There are many different centrality measures in use [15], and for this paper, we specifically looked into degree centrality (DC), betweenness centrality (BC), eigenvector centrality (EC), and closeness centrality (CC), both for FG and SG graphs. Our DC scores from both network graphs create a very different list of central nodes. The majority of the prominent nodes for both graphs are different users, and for cases when they are same, their rank is different (Table 3). These ranks are created based on how many times these nodes have retweeted or been retweeted by others in the network. As the basic difference between FG and SG is how the nodes are defined

Rank	Top Users from FG	Degree of Connections, n	Got Retweeted	Retweeted Others	Top Users from SG	Degree of Connections, n	Got Retweeted	Retweeted Others
1	TikTokDiets	583	583	0	Thinspo02468	43	0	43
2	Bonespo4eva	424	420	4	thyn_babiee	41	32	9
3	thynspoo_k	335	335	0	cicarxtxs	35	11	21
4	thin_hunnie	275	263	12	Bonespo4eva	35	33	2
5	famishedbarbie	213	204	9	restricted_vegiee	31	3	28
6	gottabth1n	189	189	0	hungri_q	30	0	30
7	Nahglossy	172	167	5	anamia_teeeenmn	29	18	11
8	EDliife01	170	169	1	lacielosweight	29	5	24
9	Skinnidaily	169	169	0	skanksubweight	29	13	16
10	bitch_n_binge	154	148	6	Eun-ice92	28	11	17

Table 3. Rank and Degree of Connections of central nodes

in terms of their connectedness, from Table 3, it is evident that the top ranked nodes from FG are primarily being retweeted by others who have less than 20 degrees of connection of their own, and thus, being downgraded in the sub-graph when those nodes are being excluded. In contrary, top nodes from SG are primarily retweeting others in the network who themselves have more than 20 connections. Additionally, we also report the average correlations among the centrality measures for better understanding of the matrices with error rate $er= 95\%$, and $\alpha = 0.05$ (Table 4). For this paper, we opted for Spearman's rank correlation coefficient (ρ) and Kendall's Tau (τ) [40] to analyze the correlations. We define three levels of correlation: high, moderate and low. The ranges of the correlation coefficients for these three levels are: high (0.75 and above); moderate (0.40 to 0.74) and low (less than 0.40) [88]. The upper section of the table highlights the rank coefficient values for each of the centrality matrices within FG and the lower section highlights values for SG . If the values between matrices are not highly correlated, they indicate distinctive measures likely to be associated with different outcomes.

FG								
Matrix	Spearman Correlation Coefficient, ρ				Kendall's Correlation Coefficient, τ			
	Degree	Eigenvector	Betweenness	Closeness	Degree	Eigenvector	Betweenness	Closeness
Degree	-	0.758	0.879	0.661	-	0.6	0.733	0.467
Eigenvector	0.758	-	0.661	0.830	0.6	-	0.511	0.689
Betweenness	0.879	0.661	-	0.624	0.733	0.511	-	0.467
Closeness	0.661	0.830	0.624	-	0.467	0.689	0.467	-

SG								
Matrix	Spearman Correlation Coefficient, ρ				Kendall's Correlation Coefficient, τ			
	Degree	Eigenvector	Betweenness	Closeness	Degree	Eigenvector	Betweenness	Closeness
Degree	-	0.720	0.757	0.369	-	0.535	0.535	0.303
Eigenvector	0.720	-	0.552	0.842	0.535	-	0.422	0.689
Betweenness	0.757	0.552	-	0.442	0.535	0.422	-	0.289
Closeness	0.369	0.842	0.442	-	0.303	0.689	0.289	-

Table 4. Spearman Correlation coefficient and Kendall's correlation coefficient for both Full Graph and Sub graph. The significant values are marked with bold ($\alpha = 0.05$)

In almost all cases, the deviations of rank correlation coefficients are almost invisible for both Spearman's (ρ) and Kendall's (τ). Looking at the significant values, where their p-value is less than α (marked as bold), all of these centrality matrices are positively correlated, and for graph *FG*, the high correlation exists within *DC* and *BC* where $\rho = 0.879$ (moderate for Kendall's $\tau = .733$). Based on these values, it can be interpreted that, for the overall network, there is a strong relation between a node's degree of connection and its probability of being the information bridge within the network. The more often a node is retweeted, the more likely that it will act as a bridge between nodes to spread information. High correlation ($\rho = 0.830$) was also observed within *EC* and *CC* for *FG* (moderate for Kendall's $\tau = 0.689$), which indicates nodes with higher influence tend to spread information faster than others within network. By imitating the same analysis on the sub-graph *SG*, we found similar correlation traits (Table 4 lower section) between the matrices where the correlation between *EC* and *CC*, and *DC* and *BC* were high (moderate for Kendall's τ) and significant. High correlation between *EC* and *BC* was also evident for both graphs, which suggests nodes with higher degree of connections not only work as bridges within the network, but also tend to be connected with other stronger nodes with similar influence. That is, popular retweeters within structured network often seek information from other active retweeters. Note that the statistics discussed above can be potentially biased due to the bias of the data we collected.

4.1.2 Sparse Community Influence by Top Users. To understand the structure of the Twitter ED network, along with nodes' centrality matrices, we also studied how these central nodes are situated within the community structure. To do that, we first looked at the clustering coefficient in the graphs we created. In graph theory, a clustering coefficient is a significant descriptive statistic that measures how closely nodes cluster together by estimating the number of linked triangles in a network [2]. In real-world networks, particularly social networks, nodes often form close-knit, densely-tied groups, with a probability that is generally greater than the chance of a connection forming randomly between two nodes [2]. For the ED retweet network, the global clustering value of the whole graph was around 0.0365 and of the sub-graph was around 0.132. A high average clustering coefficient can indicate closely connected nodes along the follower chain in a community [2]. Since the global value indicates the degree of "transitivity" of a network, higher values indicate "friends of friends" are more likely to be friends, resulting in a more clustered structure of the graph. The lower clustering value of our *FG* graph shows how sparsely the neighbors are connected in the network; however for the sub-graph, it appears to be more connected. To further our understanding of the community structure within our dataset, we also applied Louvain and Girvan Newman algorithms to identify and visualize the existence of multiple communities within this network.

Using Louvain community detection algorithms, we found 61 communities (Fig. 3a) within *FG* network graph and 14 in our sub-network graph *SG* (Fig. 3b). GN algorithm resulted in fewer communities—14 in *FG* and 6 in *SG* network graph. We have provided unique IDs for the first three biggest communities in Table 5. The modularity scores for both of these algorithms are

Louvain Algorithm							
Community ID	Members, n	Community ID	Members, n	Sub-Community ID	Members, n	Sub-Community ID	Members, n
ID FGC2	450	ID SGC3	32	ID FGSC1	227	ID SGSC1	9
ID FGC7	507	ID SGC4	34	ID FGSC2	359	ID SGSC2	6
ID FGC11	725	ID SGC8	29	ID FGSC4	89	ID SGSC3	8
Girvan-Newman Algorithm							
ID FGC2	4545	ID SGC1	174	ID FGSC2	4519	ID SGSC1	169
ID FGC12	40	ID SGC2	1	ID FGSC12	19	ID SGSC2	5
ID FGC7	32	ID SGC3	1	ID FGSC7	7	-	-

Table 5. Top 3 communities and sub-communities for *FG* and *SG* using Louvain Algorithm and Girvan-Newman Algorithm

moderate (range from -1 to 1) and indicate that the Louvain algorithm outperforms GN (for Louvain, modularity score $Q = 0.54860$; for GN, modularity score $Q = 0.45090$). By clustering communities after the first pass, the Louvain algorithm inherently considers hierarchical organization in the network. Whereas the Louvain algorithm was able to detect more communities with much more structure, the GN algorithm created one bigger community containing the majority of the nodes (for *FG*, $n = 4545$ and for *SG*, $n = 174$). Even though most of the communities created by the Louvain algorithm only include 5 or fewer nodes as community members (around 65%), the remaining 35% of communities range from 10-507 nodes. In Table 5 (left section), we have listed the 3 communities with the highest number of nodes from both algorithms. While visualizing each of these communities through graph network, we observed existence of multiple distinct clusters of nodes forming sub-communities (Fig. 3). We then ran the community detection algorithms on each of these communities again and found the existence of sub-communities within them. For example, communities FGC11 and SGC4 (the biggest communities from *FG* and *SG*), respectively had 8 (Fig. 3b, 3e) and 6 sub-communities based on Louvain algorithm. Similarly, communities FGC2 and SGC1 respectively had 5 and 2 sub-communities based on GN algorithm. The top 3 sub-communities and their node numbers are provided in Table 5 (right section).

As our community detection algorithms confirms the existence of defined communities and sub-communities within the Twitter ED network, we wanted to identify the influential nodes within these smaller clusters. Our analysis on the sub-communities FGSC1, FGSC2, FGSC4 from *FG* graph shows that for each of these communities, the central node is someone who has the highest scores in all of the centrality matrices and is consistent with our original central node list from Table 3. For example, in sub-community FGSC2, the central node is "TikTokDiets"; for sub-community FGSC1, it was "thynspoo_k" and for FGSC4, it was "Nahglossy"—all of whom belong to the original top 10 list.

As we also noticed major differences in the centrality scores between the central node and the following nodes, it can be assumed that in sub-communities, the majority of nodes are dependent on one specific node (the central node) for community connection and information flow through retweets. However, our analysis on the sub-graph's sub-communities provided different results. Even though we could identify the existence of communities and sub-communities within this smaller, more structured network, unlike the overall graph, there was no visible consistency between the central nodes based on their centrality matrices when using this same approach. The central nodes for each of these sub-communities were diverse, inconsistent, and often did not belong to the original top 10 lists. This indicates that for smaller and much more structured networks on Twitter, influence is not as static as the overall graph, but rather is dynamic and sparse within communities.

Fig. 3. Communities and Sub-communities on Twitter Using Louvain Algorithm on Full Graph

4.2 Validating Contents within ED Community Online

As we investigated the concept of influence within ED network through analysis of prominent nodes, in this section we also explored what prominent contents, such as hashtags and media, were shared within the communities.

4.2.1 Assessment of Hashtags. While our inductive analysis provided understanding on established ED hashtags that are prominent within the Twitter ED network, we also identified the prevalence of nuanced and newly emerging hashtags with unique linguistic meanings. A total of 75,676 hashtags were attached to the 32,000+ posts in our dataset, of which 1,141 were unique. On average, there were 4 tags attached to each post ($SD = 2.0$; range 1-14). Apart from the three primary hashtags that we started working with, hashtags like #meanspo, #ed, #sweetspo, #bonespo, #anorexia, and #bulimia were some of the most used hashtags in our dataset (Fig. 4).

Fig. 4. Top 10 hashtags within the dataset

The categories that evolved for this part of the analysis were derived from the hashtags associated with each of the posts in our dataset. We created a small classification codebook for the hashtags;

Table 6 highlights and defines the primary categories. While the majority of these categories have been consistently perceived through previous literature, we did observe new emerging categories of hashtags, such as meanspo and trends/viral platforms, that have not been properly introduced and discussed. Even though the hashtag category Viral Platforms and Trends only encompasses 0.2% of the total hashtags observed within our dataset, it signifies emergence of new social media platforms within the Twitter ecosystem through hashtagging. Though some of these platforms (e.g., Instagram, Tumblr) have been popular for years, platforms like Tiktok, Douyin, and OnlyFans are comparatively new, have been trending/viral within younger generations, and are making an impact within the Twitter ED network as well.

Categories	Details	Tag Examples
Inspiration (27.6%)	Terms associated with ED-specific support and other forms of empowerment. Includes inspiration related to pop culture, race/ethnicity, and gender	#thinspo, #poethinspo, #kpopthinspo
General Eating Disorders (ED) (29%)	Terms associated with general tweets related to eating disorders	#EDtwit, #edthings, #eatingdisorder
Anorexia (19.2%)	All terms associated specifically with Anorexia are captured in this category	#anorexia, #proana, #anorexia
Bulimia (3.03%)	All terms associated specifically with Bulimia are captured in this category. Unlike Anorexia, this category also includes central behaviors and activities crucial to the disease of bulimia like bingeing and purging.	#bulimia, #promia, #mia
Body Part (2.27%)	This category encompasses all terms associated with anatomical parts of the human body. It also includes terms related to desire to see bones, such as bonespo.	#legspo, #ribcage, #bonespo
Body Image (4.36%)	This category encompasses all terms associated with the physical state or body image of the individuals, such as fat, skinny, thin, etc.	#skinny, #slimspo, #FlatTummy
Coach and Tips (0.48%)	All terms related to finding coaching and tips for ED motivations	#coaching, #proanacoach, #proanatips
Emotional/Mental State (0.1%)	Emotional terms associated with sadness, depression, anxiety, or behaviors associated with these sentiments. These sentiments could be reflective of oneself, the community, or the world.	#depressed, #anxiety, #mentalillness
Diet/Food (0.25%)	Food- and beverage-related terminology, as well as diet-related terms and charts are included here	#fasting, #safefood, #abcdiet
Grunge/Goth (0.05%)	This category encompasses gothic and grunge terms associated with eating disorder lifestyle	#grunge, #goth, #darkgrunge
Meanspo (4.54%)	This category encompasses only the terms specifically related to the "mean" way of inspiration. These inspiration/motivation may be reflective of oneself, the community, or the world.	#meanspo, #mean, #Meanspo
Suicide/Self Harm (0.12%)	Explicit and implicit suicidal or self-harm related ideation are included in this category	#proselfharm, #selfharm, #suicide
Viral Platforms (0.2%)	This category includes terms individuals used to connect their ED motivations from other social media platforms. Platforms like Tiktok, Douyin, Instagram, Tumblr , Onlyfans are included here	#douyin, #tiktok, #onlyfansgirls

Table 6. Hashtag categories

4.2.2 New observed activity: #Meanspo. Of the total hashtags within our dataset, 4.54% were related to the emerging pro-ED trend of Meanspo, including 2837 instances of the specific hashtag #meanspo. Our data shows that meanspo, short for "mean inspiration", consists of overly critical and insulting statements hurled at those who are "not thin enough," (either oneself or others) intended to motivate them to stop eating, purge, and lose weight. While previous literature has discussed positive and negative thinspirational posts in pro-ED online communities [107], the #meanspo hashtag has never been specifically investigated, and many of the meanspo tagged tweets go beyond this general notion of negative inspiration and become overly harassing at times. This paper uses #meanspo to classify these negative inspirations for a better understanding of not only the negative contents shared and sought out on ED Twitter, but also the extremely harassing or toxic environment that frequently lurks behind general/trivial terms like "negative inspiration". Of note, although other tweets may have included content that was mean or harsh in nature, we only categorized tweets where users directly tagged the post with a variation of the "meanspo" hashtag.

One class of #meanspo tagged tweets focused on support through *disapproval or negative contexts* where a user tries to "motivate" oneself or others through mean or passive-aggressive disapproval of their eating habits or related lifestyle choices. Examples of this type of post include:

- Restrict..or you are going to regret it
- If you really wanted to be skinni would you eat that?
- Just bcz they don't call you fat on ur face mean you are skinny..try harder

Media Type	General Posts (n=510)	Posts with No Text (n=200)
Photo	47.6 %	46.7%
Video	20.2%	17.4%
Gifs	1.0%	0.0%
Unavailable/Suspended	31.2%	35.9%

Table 7. Details on Media Type

Another form of #meanspo tagged tweets were more *antagonistic or aggressive* in nature and were extremely derogatory towards oneself or others. Extreme bullying and harassing contents were very common within this category and often included extreme fat-shaming and name-calling as "inspiration". Examples include:

- I'm a worthless fat pig and nothing can stop me until I die
- Stop lying to yourself whore! You fucking loser, you're never going to lose weight!
- You think you can eat that and loose weight you disgusting fucking fat piece of shit?

Finally, there were people directly *soliciting meanspo* to be directed at them or joining a group that would support this type of negative pro-ED support. Users would actively ask to join or create groups or group chats that were specifically focused on sharing/providing meanspo. Some users were specifically seeking a meanspo "coach" or "buddy" with experience motivating others to lose weight in this manner. Examples include:

- Anyone interested in a meanspo gc? Looking to send/receive with like minded people
- where are the proanna coaches or meanspo lovers? looking for a new texting buddy?
- I desperately need a meanspo gc where we can talk shit abt each other
- I am in need of a meanspo bae

These categorizations of #meanspo tag highlight the use of negativity as a driving force for people to comply with the difficulties of their ED journey. While our analysis suggests new emerging activities through introducing new hashtags within online ED communities, exploration of such tags prompts deeper discussions on the bullying and self-harassment that is integrated within the ED network.

4.2.3 Media Analysis. Along with hashtags, we also looked at different media attached to the tweets in our dataset and identified certain ED-related archetypes. Out of around 8,000 media, we randomly selected 710 (around 9%) and manually coded them. We separated our media links into posts that included text in addition to the media link (n=510) and posts with no additional text other than hashtags (n=200). Our initial analysis showed that, within the first category, around 47.6% of the posts were photos, 20.2% were videos, 1% were gifs, and the rest of the media were either suspended, unavailable, or access limited (Table 7). We observed similar presentation of media within the second category. We categorized media into archetypes that include, but are not limited to, Thinspiration, Celebrity Idealization, Trends and Viral Platforms, Updates on ED Journey, ED-Related Memes, Resource/Information Sharing, and Unrelated Images. While all media did not fall into one of these foci, the following are representative of a majority of posts within our dataset.

General Thinspiration: Media in this category offer inspiration to both oneself and other users within the network. Thinspiration or "thinspo" are media, such as images of severely skinny body parts, that are intended to encourage others to strive for thinness as the "ideal" body type. Fig. 5 is an example of thinspiration. This category also includes "bonespiration" or "bonespo," characterized by the ability or desire to see as many of your bones through your skin as possible. Some media in this category also include goth- or grunge-inspired ED contents, where the focus is on different

Fig. 5. Thinspo & Celebrity Idealization

Fig. 6. TikTok Videos on ED Lifestyle

Fig. 7. ED Journey Updates

Fig. 8. Resource and Information Sharing

body parts featuring dark or black clothing or overall aesthetic.

Celebrity Idealization: This category includes thinspiration related to celebrities whose bodies and thinness are idealized, like Kendall Jenner, Ariana Grande, and various K-pop stars, such as Itzy Lia (Fig. 5). Photos and videos related to these and other celebrities were commonly used as pro-ana and ED lifestyle motivation in our dataset.

Connections to Other Social Platforms: Media in this category promote similar types of thinspiration, with the additional criteria that they were initially posted on other social media platforms (e.g., the Chinese app Douyin and its international version, TikTok) prior to being shared on Twitter (Fig. 6). Previously seen in our hashtag analysis, TikTok and Douyin are popular social media platforms upon which millions of users create a wide variety of short-form videos. In our dataset, many of these videos include individuals (not necessarily the Twitter user) posing or dancing in front of camera while showing off their very thin body, clothes, or specific body parts. Fig. 6 includes examples of such videos.

Updates on the ED Journey: Media in this category focus on the journey and updates associated with weight loss. Fig. 6 and 7 provide example of these type of images, including current personal before-and-after shots. Some users highlighted this journey through screenshots from different mobile apps that store information on calorie intake, water intake, amount of weight loss, exercise timeline, and more.

Resource/Information Sharing: This category includes media sharing resources and information related to the ED lifestyle, like pictures of foods suggested and/or consumed by users (often with calorie counts included) as part of diet regime, type of exercises for losing weight (again, sometimes with calorie counts), and information on TV shows, movies, or books related to ED (Fig. 8).

4.2.4 Validation by Central Nodes within ED Network. We wanted to validate our analysis of contents through the sharing pattern of the prominent nodes within the ED network and exploring

the contents of the hashtags they used and media they shared. Our preliminary analysis on the overall data showed that the majority of the prominent nodes from FG shared actual contents, rather than retweets, as opposed to the prominent nodes from the sub-graph, where the tendency of retweeting was much more prominent (Fig. 9).

Fig. 9. Actual vs. Retweets Frequencies of Prominent Nodes

First, we compared the total number of individual hashtags with the total hashtags shared by these nodes (see Table 8). While the frequency of hashtags related to #thinspiration, #edtwitter, and #ana/#proana was very high within these nodes (as expected, as they were our base hashtags), we also observed additional prominent hashtags, such as #edmeme, #bodygoal, #skinny, and #meanspo. Particularly, we are interested in #meanspo, as it suggests that antagonistic or harassing behaviors among ED community members are affected by these influential nodes within the networks. We also noticed consistency between key nodes and their media sharing pattern in the ED network (see Table 8). Similar to our overall finding, the majority of their shared media were based on Thinspiration, Resource Sharing, Tiktok and Douyin Thinspiration, ED-Related Memes, and so on. These results reconfirm our inductive analysis of the contents shared within the ED network and suggest possible influence of the central nodes on what contents are widely shared and become popular. However, it is worth noting that roughly 32.5% of the contents in our randomly selected sample had been removed by either the user (i.e., account or tweet deletion) or Twitter (i.e., account suspension). Although these contents were available at the time of data collection, their subsequent removal or deletion hampered our content analysis. This censorship was apparent within our dataset's major nodes—multiple entire accounts had been deleted or suspended, rendering their content inaccessible.

4.3 Identifying Emerging ED-Related Linguistic Indicators Based on Topic Modeling

While the previous section identified popular contents through the analysis of hashtags and media links, in this section, we analyzed the actual text of the tweets and identified emerging topics of discussion within our dataset. After data pre-processing, we obtained 32,530 tweets to feed into our topic modeling analysis, where we used both LDA and NMF algorithms to identify the prominent topics. These tweets contained 228,357 tokens with a vocabulary size of 9,316. The mean length of the tweets was 5.5 and the max length was 18.

4.3.1 Analysis of Topics. When specifying a topic model, the number of topics (K) must be defined. To do this, we ran several candidate models with varying numbers of topics, and for $k=10$ topics, we received the highest coherence and perplexity scores for both LDA ($CS=0.34876$, $perplexity=-5.843$) and NMF ($CS=0.38826$) models.

Central Node	Prominent Use of Hashtags (%)	Prevalent Type of Content Shared
TikTokDiets	#Thinspo/#thinspiration (100%)	Unavailable: Could not check
Bonespo4eva	#thinspo (80%), #edtw (80%), #meanspo (58%), #edmeme (10%)	ED memes, thinspo, Tiktok, resources
thin_hunnie	#thinspo (73%), #proana (50%), #meanspo (58%)	Unavailable: Could not check
thynspoo_k	#thinspo (100%), #skinny (90%)	thinspo, update, Tiktok
famishedbarbie	#edtw (58%), #thinspo (40%), #proana (35%), #anorexia (32%), #meanspo (6)	Unavailable: Could not check
Nahglossy	#thinspo (100%), #meanspo (18%)	Unavailable: Could not check
Thinspo02468	#thinspo (76%), #ana (41%), #meanspo (14%)	thinpo, Tiktok
restricted_vegie	#thinspo (80%), #bodygoal (24%), #meanspo (7%)	ED memes, thinspo
hungri_q	#thinspo (57%), #proana (34), #meanspo (10%)	thinspo, grunge, not me
anamia_teeeen	#thinspo (59%), #edtw (68%), #proana (34%)	thinspo, resources, food, ED meme, Tiktok
EDlife01	#thinspo (43%), #proana (29%), #meanspo (14%)	Thinspo, Tiktok, Resource
Skinndaily	#edtw (66%), #bslyw (39%), #thinspo (32%)	thinspo
bitch_n_bingee	#edtw (74%), #thinspo (42%), #proana (42%)	thinspo, resources
skanksubweight	#meanspo (50%), #thinspo (21%), #edtw (14%)	thinspo, resources

Table 8. Content shared by central nodes: hashtags (percentage) and content archetype

Fig. 10. Topic Reliability between LDA and NMF

We manually analyzed the top keywords and most relevant tweets generated by the models for each topic, and came up with 10 primary themes for both LDA and NMF. Additionally, for each of these primary themes, we identified multiple secondary themes that further explain the primary themes. We have added the primary and secondary themes for both models in the appendix. To measure the reliability of our topic models, we compared the two models to identify similarities in resulting keywords under each topic. The results of this comparison are given in Fig. 10, where the column numbers represent LDA topics and the row numbers represent NMF topics. The rate of convergences between different topics for both models range from 0-60%.

4.3.2 Qualitative Assessment of Topics. The topic categories resulting from the qualitative analysis of the LDA and NMF topic models highlight standard content within our dataset (Table 9). The topics with similar themes were grouped into 5 categories: ED Inspiration and Motivation, Weight/Fitness, Community Building and Maintenance, Engaging the Community, and Reflections and Desires. Below, we unpack several topics that presented more nuance or new aspects of known ED-related activities.

Topic Categories	LDA Topics	NMF Topics
ED inspiration and motivation	-Specific ED Inspiration -Thought processes during active ED -General presentation of ED	-Support/Motivation during active ED -ED focused aspirations
Weight/Fitness	-Weight loss setting/status -Fitness/diet	-Weight loss
Community Building and Maintenance	-Community building/Maintenance	-Community building -Community-based characteristics and curating -Popularity of major community member
Engaging the Community	-Engagement with sub-communities -Best Practices and plans -ED experiences and values	-Disorder-specific engagement -Targeted engagements and presentations of ED journey
Reflections and desires	-Personal reflection on body image	-Desired ED lifestyle -ED-specific reflection

Table 9. Topic Categories Created from LDA and NMF

ED Inspiration and Motivation: This topic included traditional forms of ED-related inspiration and motivation. Inspiration focused on the ideal body type or activities to achieve it, and included pro-ED support both positive and negative in tone, including the aforementioned "meanspo.". The concept of negativity as a form of inspiration or support in pro-ED behaviors is not novel, yet the emergence of communities forming around and seeking out this particularly direct, antagonistic, and sometimes severe harassment has not yet been explored. Within this category, we also saw posters connect traditional sub-communities (#proana, #thinspo, #edtw) with both gender (#male-spo) and ethnicity (#pochinspo). Within the media content of these posts, users shared traditional ED media (e.g., specific body parts, diets, the curation of idealized celebrity bodies/lifestyles, and before-and-after/progress), yet additionally made them gender- or ethnicity-specific.

Community Building and Maintenance: Topics within this category centered around how members actively or passively carried out these activities. A prominent aspect of this data is the amount of censorship or moderation. While there were high levels of moderation within the Twitter platform (see Table 8), we also saw examples of censorship through gatekeeping within the communities by members who were actively opposed to cross-platform integration of the TikTok/Douyin community within the ED Twitter space. Some posts explicitly dissuaded new users from joining or participating in ED Twitter if they were an active TikTok user, as well as from re-posting TikTok videos. This theme also included examples of community members actively seeking to build specific group chats or find support, often centered around a specific theme or activity (e.g., meanspo, accountability in diet/exercise, and sharing of best practices). Additionally, the high profile or influence of specific members of the community is reflected in this topic. For example, an entire NMF topic was dedicated to the community retweeting a specific prominent node. Unfortunately, that data could not be analyzed, as that node's account was suspended by the time qualitative analysis took place.

Reflections and Desires: This category includes themes related to users' explicit desire associated with various aspects of an ED (e.g., ideal body, ability to restrict, the way others perceive them) as well as emotions associated with these aspects, reflections on where they are in their ED journey, and the challenges/rewards of the ED journey. These tweets reflected a full spectrum of emotional and mental states, ranging from negative (e.g., sadness, frustration, defeat, self-hate) to positive (e.g., pride, happiness, optimism). However, in general, there was a negative skew, especially in

posts where users indicated their desperation or longing (want, wish, kill) for certain ED lifestyle and body traits (skinny, thigh gap, visible bones).

5 DISCUSSION

5.1 Influential Users and Exposure to Unhealthy Lifestyle

Interactivity among online communities is an important factor impacting how members' attitudes, intentions, and perceptions regarding health-related behaviors are influenced [70]. Identifying the influential nodes in our dataset and their sharing pattern of contents provides valuable information on the ED network's node-level structure. The large difference in number of nodes and the rank difference between the retweet network (FG) and the subset of this network (SG) highlight two important implications: first, a majority of the general nodes in FG mainly retweet the central nodes and do not frequently interact with each other (thus, have fewer degrees), and second, the central nodes from FG are primarily being retweeted, rather than retweeting others in the ED network and are the main content creators of the network. We explain these implications further below.

5.1.1 Homophily and Social Influence in Retweet Network. As this study aimed to identify influential individuals within the ED network, we observed certain disparities between the ranks of popular nodes (between FG and SG) due to the diverse retweeting behaviors within users. According to social network structure theory, the degree of connectivity of a node in a Twitter network can be a measure of a user's popularity and social influence [49, 84]. While certain nodes had a higher degree of connections within the ED network due to being frequently retweeted by others, the sub-graph created from this network had much more inter-group interaction where the central nodes not only were retweeted, but also retweeted other stronger nodes (Table 3). These patterns of communication are not uncommon within structured online health communities where ties between nodes are stronger, as opposed to a more unstructured network [112].

This communication pattern within structured network can be supported by the idea of homophily, which asserts that users in a social system are more likely to bond with those who are "similar" to them than those who are "dissimilar" [45]. Research on food blogging networks found that bloggers tend to interact with other bloggers that create or value similar content in order to maximize their exposure to interested audiences (e.g., through cross-promotion) or gain credibility through ties to well-known bloggers in the same subcategory [6]. While we observed homogeneous ties within the ED network, where users retweeted nodes with similar interests, it was more evident and frequent within the sub-graph network. This retweeting behavior of nodes highlights a certain disparity between content production and engagement/consumption of content (e.g., retweets): the central nodes from the FG network mostly create tweets rather than retweet others, as opposed to central nodes from the sub-graph (see Fig. 9). While mutual connections between nodes are typically analyzed within homogenous networks, these connections can also be one-directional, with general nodes seeking guidance or information from specific influential nodes [120].

Communication structure (who communicates with whom) is key for the study of peer influence on health behaviors [93, 111]. For example, Twitter networks of diabetes and diets [55] highlight how certain users are seen as opinion leaders who affect other users' attitudes and behavior [112, 151]. Though individuals in such networks may not actively participate by providing their own content, they are still able to derive value and potentially shape ideas and norms by being influenced through these central nodes [63]. As the ED network on Twitter includes multiple communities and sub-communities, such social influence can be perceived through community-level interaction between central and general nodes. Even though studies like [63] on drug abuse networks suggest that cascades have significant structural virality (i.e., they are driven by numerous significant nodes), for the ED network, we observed it is also possible to have smaller communities

that are led by single and specific central nodes who mostly create contents and are responsible for influencing others (e.g., TikTokDiets). While influence within sub-communities led by multiple nodes is difficult to control, being able to identify these prominent nodes could provide a potential opportunity for platforms to implement regulation against sharing risky ED contents.

5.1.2 Exposure and Unhealthy Adoption of ED Lifestyle. Our findings on influential nodes within our dataset support our content analysis findings, which showed that popular contents within the network are shared by its prominent users. Although previous literature has explored popular contents within online ED communities [30, 107], this study connects those contents with node-level participation. As most interactions within our dataset occurred through retweeting central nodes, this higher exposure to the unhealthy behaviors promoted by these contents may influence higher adoption of such behaviors by other users.

Previous research suggests that homophilous relationships within a community can enhance the dissemination of behavior among individuals [26, 45], and information that is sourced from or endorsed by a prominent figure in a social network may have increased value to other users [111]. Similarly, individuals who work as bridges between multiple nodes and groups, like our prominent nodes with higher *BC*, further enhance this diffusion of information. Research also demonstrates that exposure to information that promotes the thin ideal encourages self-objectification in individuals [60, 74], which can lead to disordered eating behaviors. This increased exposure to negative social influence and harmful information from important nodes within a network raises the likelihood of adoption among general users, particularly those with health concerns.

5.2 Antagonistic Motivation for ED Success

5.2.1 Motivation vs. Harassment. Negativity is a common motivating factor within the ED community [107, 143], including the emergence of meanspo—extremely aggressive, harsh, or abusive comments directed toward self or others in an effort to promote disordered eating behaviors [3]. Meanspo was prevalent within our dataset and included a unique phenomenon in which users specifically requested this form of negativity and discussed creating group chats exclusively to "support" each other in this way. This use of personal insults that border on emotional abuse are likely to have an impact on the clinical manifestations of eating disorders, as negative feedback has been implicated in the development of increased body dissatisfaction [23, 46] and bulimic behaviors [121]. Research showed that individuals who reported a greater negative feedback seeking style via Facebook and received a high number of comments on these posts were more likely to report higher eating restraint [68] and body dissatisfaction at a later time [23, 46]. This elicitation of negative feedback provides a type of support that is needed within this community to continue with harmful, pro-disease activities.

While such antagonistic communications are masked as motivation/inspiration, and were sometimes even specifically sought out by the users in our dataset, these comments often blur the line between motivation and harassment. Online harassment is an issue of public concern and discussion, particularly for vulnerable populations, in the context of persistent problems like cyberbullying [41, 97, 98] and hate speech [50]. Lenhart et al. [78] argue that "online harassment is defined less by the specific behavior than its intended effect on and the way it is experienced by its target". Prior work has indicated that social media sites (e.g., Reddit, Twitter) can provide a false sense of anonymity, lowering users' social inhibition and encouraging more aggressive communication [118]. As such, some users may perceive their online conduct as harmless or acceptable within the norms of a given community while others may define it as harassment.

While harassment is frequently defined as an exterior act directed at another person [149, 152], it may also be internalized, with a person harassing themselves through self-injurious behaviors [61],

which reflected online is known as digital self-harm or self-harassment/cyberbullying [103, 108]. According to Pater and Mynatt [104], harassment within online ED spaces can include speech and conduct that leads to or reinforces non-suicidal, yet purposeful, injury or damage of an individual's physical well-being. Many tweets in our dataset contained extremely negative contents influencing others or oneself toward intentional harm via promotion of ED-related behaviors. While these act of negativity may count as motivations within online ED networks, they reinforce severe and dangerous lifestyles within users.

5.2.2 Cyber-disinhibition and Group Conformity. The relative freedom to participate in hostile and provocative deviant activities with the goal of offending and hurting others is common within online communities [35, 82], as also seen within ED networks. Such de-individuation occurs when users believe they are not being individually assessed, allowing themselves to behave in ways they would not otherwise. Such conduct is referred to as cyber-disinhibition or the toxic online disinhibition effect [153], which embodies vulgar language, harsh remarks, hate speech, and even threats that would be highly unusual in a face-to-face context [141]. This can be seen in pro-ED blogs [21], social loafing in online communities [126], hate sites [33], cyberbullying [67] and more. Disinhibition is often amplified if the user has a sense of anonymity [98], which is common in online ED spaces, where accounts often use generic ED-related usernames that do not reveal the individual's identity, but rather display their conformity with community values. Even the content shared is often generic and of other people, which potentially may make individuals feel less inhibited when using antagonistic and severely harsh language.

Feelings of de-individuation within ED networks can come from users' experiences with group conformity where they lose their personal identity and connect with the group identity. As the Twitter ED network has strong evidence of communities and sub-communities that exacerbate harmful behaviors within users through antagonistic motivations, it is possible for community members to take on a de-individualized group identity and follow the group norms established by influential members. Such conformity can work as an uncontrollable external condition that directly and indirectly influences users to seek and provide extremely antagonistic tweets as motivations, and may socialize them into a normalized culture of harassment. Therefore, it is essential that we learn these internal group norms and incorporate discussion of such motivation into the context of one's self-harm behaviors online.

5.3 Censorship within ED Communities

Censorship within online ED spaces is a phenomenon that has garnered interest in the HCI community [29, 30, 108]. A substantial number of tweets and accounts in our dataset had been removed between the time of data collection and data analysis, by both users and the Twitter platform. The latter serves as a potential example of censorship that occurs due to self-expression of ED behaviors and activities online, as this can go against the guiding policies (e.g., Terms of Service, Community Standards) of many online platforms, which characterize EDs in varying ways, including harassment, harm, self-harm, self-injury, abuse, and even explicitly ED-specific guidelines [108]. Censorship through in-community gatekeeping to maintain community norms was also evident in the dataset. While the cross-posting of ED content from one platform to another is common [59, 107], some Twitter users in our dataset were vocally opposed to the integration of TikTok users and content. This insider vs. outsider conflict points to a set of social norms that may be difficult to perceive for those on the periphery, but are ingrained within those at the core of the community [77]. The internal community norms and values of online ED communities often run afowl of broader platform community standards [108] or smaller sub-communities rules [58]. The

conflict of these standards has ramifications not just for the individual, as seen by the personal effects of censorship [122], but also for the community as a whole and how it organizes itself [29].

Censorship also impacted our research process. As censorship was frequently used to restrict activities/contents within Twitter ED networks (as noted in the results), it potentially hampered and limited our qualitative assessment of media links. Past research has shown the importance of assessing the media of ED content, as it can often paint a more nuanced or different story than the text itself [107]. Thus, we should create space for conversations and further research related to the general community harm caused by pro-ED content versus the harm to sub-communities when expressions of their illness are censored.

6 LIMITATIONS AND FUTURE WORK

There were several limitations to this study. First, the authors are not members of the ED community on Twitter. As outsiders, it is possible to miss latent signals or misinterpret aspects of the ED activities uncovered in qualitative coding. To address this limitation, we consulted with our clinical partners at [redacted] whenever we had questions during the coding process. Second, we used the NMF and LDA algorithms for the analysis, which are both unstructured. A point of future work could be to analyze such a dataset with both structured and unstructured algorithms and explore any clinically and topic relevant differences between the two approaches. Third, the initial search terms we used were general in nature. Prior work has found differences across gender and ethnicity regarding the language used to engage in online ED communities [44, 81, 110]. As, we did see traces of gender and ethnic diversity in our dataset (e.g., #menwithED and #POCthinspo), exploring the differences in online ED community structures related to gender and ethnicity is another area deserving of further research. Finally, this analysis only included tweets spanning a 30-day period. Looking at data over a longer duration of time may provide further or more nuanced insight. Future work is also needed to understand the interplay of node-level activities *across* platforms, such as Twitter, TikTok/Douyin, Instagram, and YouTube, as observed within our data.

Translating findings from online ED community research into clinical practice is another important future direction. Providers have said that they believe social networks have a negative impact on patients and that addressing social media activity in the clinical setting should be explored [105]. As providers lack knowledge in this domain [105], creating feedback loops would be a positive next step towards the meaningful integration into clinical workflows and practice. For example, patterns of influence (content or individuals) could help providers target therapeutic approaches, understand triggers and responses, and keep them abreast of the technical affordances/trends that could have implicit or explicit impacts on their patients. Connecting the technical with the clinical will take more than just surface-level integration, but true embedded partnerships that will allow for a more meaningful and deeper engagement and facilitate the translation across domains. While we have used the lens of eating disorders for this research, we believe that applying this methodology to other health domains (e.g. other chronic illnesses) would potentially illuminate new aspects of their online presentation.

7 CONCLUSION

This study analyzes the ED network on Twitter, based on a dataset collected from a specific timeline, using both qualitative and quantitative measures. As social media can have dangerous impacts on ED patients, this study aims to understand how ED influences are shaped on Twitter, from network structure as well as from content adoption. Details from this study will not only help researchers and clinicians to identify possible sources of influences for ED patients through social media participation, but also will help policymakers to structure appropriate preventive methods to limit promotion of such harmful behaviors online. Insights from our analysis on node-level

influence identify the importance of understanding the nuanced community structure within online ED networks, as the roles of the central nodes and their influence over other nodes vary depending on how these smaller communities and sub-communities are formed. Additionally, insights from our content analysis can help researchers and clinicians to identify particular online channels and technological affordances that foster adverse effects on their patients, as well as discover what common and emerging behaviors and environmental factors are contained within these platforms in relation to eating. The contributions of this study also illustrate the importance of additional studies in the social computing area that not only investigate the effects of social media on population-level ED patients, but also at the person level. Studying these impacts is a crucial component in linking HCI research to clinical practice in this area, and eventually learning how to clinically treat digital self-harm in the future.

REFERENCES

- [1] [n.d.]. Twitter premium APIs – Twitter Developers | Twitter Developer. <https://developer.twitter.com/en/products/twitter-api/premium-apis>
- [2] 2018. Clustering Coefficient in Graph Theory. <https://www.geeksforgeeks.org/clustering-coefficient-graph-theory/>
- [3] 2020, May 21. *What you need to know about the dangers of Meanspo*. <https://info.axis.org/blog/what-you-need-to-know-about-the-dangers-of-meanspo>
- [4] 2021. Prevalence, Onset and Course of Eating Disorders. <https://www.mentalhelp.net/eating-disorders/prevalence-onset-and-course/>
- [5] Youcef Abdelsadek, Kamel Chelghoum, Francine Herrmann, Imed Kacem, and Benoît Otjacques. 2018. Community extraction and visualization in social networks applied to Twitter. *Information Sciences* 424 (2018), 204–223.
- [6] Liu H.-Tang L. Philip S. Y. Agarwal, N. 2012. Modeling blogger influence in a community. *Social Network Analysis and Mining* 2, 2 (2012), 139–162. <https://doi.org/10.1007/s13278-011-0039-3>
- [7] Yong-Yeol Ahn, Seungyeop Han, Haewoon Kwak, Sue Moon, and Hawoong Jeong. 2007. Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th international conference on World Wide Web*. 835–844.
- [8] Amaal Alruwaily, Chelsea Mangold, Tenay Greene, Josh Arshonsky, Omni Cassidy, Jennifer L Pomeranz, and Marie Bragg. 2020. Child social media influencers and unhealthy food product placement. *Pediatrics* 146, 5 (2020).
- [9] Monica Anderson and JingJing Jiang. 2018, May 31. Teens, Social Media Technology 2018. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>
- [10] Jon Arcelus, Alex J Mitchell, Jackie Wales, and Søren Nielsen. 2011. Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Archives of general psychiatry* 68, 7 (2011), 724–731.
- [11] Alina Arseniev-Koehler, Hedwig Lee, Tyler McCormick, and Megan A Moreno. 2016. #Proana: Pro-eating disorder socialization on Twitter. *Journal of Adolescent Health* 58, 6 (2016), 659–664. <https://doi.org/10.1016/j.jadohealth.2016.02.012>
- [12] Brook Auxier and Monica Anderson. 2021, April 07. Social Media Use in 2021. <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>
- [13] Eytan Bakshy, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. 2011. Everyone’s an Influencer: Quantifying Influence on Twitter. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining* (Hong Kong, China) (WSDM ’11). Association for Computing Machinery, New York, NY, USA, 65–74. <https://doi.org/10.1145/1935826.1935845>
- [14] Anna M Bardone-Cone and Kamila M Cass. 2007. What does viewing a pro-anorexia website do? An experimental examination of website exposure and moderating effects. *International Journal of Eating Disorders* 40, 6 (2007), 537–548.
- [15] Michele Benzi and Christine Klymko. 2014. A matrix analysis of different centrality measures. *arXiv preprint arXiv:1312.6722* (2014).
- [16] Lobban F.-Belousov M. Emsley R. Nenadic G. Bucci S. Berry, N. 2017. #WhyWeTweetMH: understanding why people use Twitter to discuss mental health problems. *Journal of medical Internet research* 19, 4 (2017), e107. <https://doi.org/10.2196/jmir.6173>
- [17] Lindsay Blackwell, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. 2017. Classification and Its Consequences for Online Harassment: Design Insights from HeartMob. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 24 (Dec. 2017), 19 pages. <https://doi.org/10.1145/3134659>
- [18] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.

- [19] Pascoe-C. J. Boero, N. 2012. Pro-anorexia communities and online interaction: Bringing the pro-ana body online. *Body & Society* 18, 2 (2012), 27–57. <https://doi.org/10.1177/1357034X12440827>
- [20] Schenk-S. Wilson J. L. Peebles R. Borzekowski, D. L. 2010. e-Ana and e-Mia: A content analysis of pro-eating disorder web sites. *American journal of public health* 100, 8 (2010), 1526–1534. <https://doi.org/10.2105/AJPH.2009.172700>
- [21] Sarah R. Brotsky and David Giles. 2007. Inside the “Pro-ana” Community: A Covert Online Participant Observation. *Eating Disorders* 15, 2 (2007), 93–109. <https://doi.org/10.1080/10640260701190600>
- [22] Palmer S.-Togher L. Hemsley B. Brunner, M. 2019. ‘I kind of figured it out’: the views and experiences of people with traumatic brain injury (TBI) in using social media—self-determination for participation and inclusion online. *International journal of language communication disorders* 54, 2 (2019), 221–233. <https://doi.org/10.1111/1460-6984.12405>
- [23] Kwitowski-M. A. Mazzeo S. E. Burnette, C. B. 2017. “I don’t need people to tell me I’m pretty on social media:” A qualitative study of social media and body image in early adolescent girls. *Body Image* 23 (2017), 114–125. <https://doi.org/10.1016/j.bodyim.2017.09.001>
- [24] Gabriella Casalino, Ciro Castiello, Nicoletta Del Buono, and Corrado Mencar. 2018. A framework for intelligent Twitter data analysis with non-negative matrix factorization. *International Journal of Web Information Systems* (2018).
- [25] Damon Centola. 2010. The spread of behavior in an online social network experiment. *science* 329, 5996 (2010), 1194–1197.
- [26] Damon Centola. 2011. An experimental study of homophily in the adoption of health behavior. *Science* 334, 6060 (2011), 1269–1272.
- [27] Stevie Chancellor, Yannis Kalantidis, Jessica A. Pater, Munmun De Choudhury, and David A. Shamma. 2017. Multimodal Classification of Moderated Online Pro-Eating Disorder Content. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI ’17). Association for Computing Machinery, New York, NY, USA, 3213–3226. <https://doi.org/10.1145/3025453.3025985>
- [28] Stevie Chancellor, Zhiyuan Lin, Erica L. Goodman, Stephanie Zerwas, and Munmun De Choudhury. 2016. Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work Social Computing* (San Francisco, California, USA) (CSCW ’16). Association for Computing Machinery, New York, NY, USA, 1171–1184. <https://doi.org/10.1145/2818048.2819973>
- [29] Stevie Chancellor, Zhiyuan (Jerry) Lin, and Munmun De Choudhury. 2016. “This Post Will Just Get Taken Down”: Characterizing Removed Pro-Eating Disorder Social Media Content. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI ’16). Association for Computing Machinery, New York, NY, USA, 1157–1162. <https://doi.org/10.1145/2858036.2858248>
- [30] Stevie Chancellor, Jessica Annette Pater, Trustin Clear, Eric Gilbert, and Munmun De Choudhury. 2016. thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work Social Computing* (San Francisco, California, USA) (CSCW ’16). Association for Computing Machinery, New York, NY, USA, 1201–1213. <https://doi.org/10.1145/2818048.2819963>
- [31] Chiu-C. H. Miao N. F. Chen P. H. Lee C. M. Chiang J. T. Chang, F. C. 2016. Predictors of unwanted exposure to online pornography and online sexual solicitation of youth. *Journal of health psychology* 21, 6 (2016), 1107–1118. <https://doi.org/10.1177/1359105314546775>
- [32] Jonathan Chang, Jordan Boyd-Graber, Chong Wang, Sean Gerrish, and David M Blei. 2009. Reading tea leaves: How humans interpret topic models. In *Neural information processing systems*, Vol. 22. Citeseer, 288–296.
- [33] Xu-J. Chau, M. 2007. Mining communities and their relationships in blogs: A study of online hate groups. *International Journal of Human-Computer Studies* 65, 1 (2007), 57–70. <https://doi.org/10.1016/j.ijhcs.2006.08.009>
- [34] Jilin Chen, Werner Geyer, Casey Dugan, Michael Muller, and Ido Guy. 2009. Make New Friends, but Keep the Old: Recommending People on Social Networking Sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (CHI ’09). Association for Computing Machinery, New York, NY, USA, 201–210. <https://doi.org/10.1145/1518701.1518735>
- [35] Christy MK Cheung, Randy Yee Man Wong, and Tommy KH Chan. 2020. Online disinhibition: conceptualization, measurement, and implications for online deviant behavior. *Industrial Management & Data Systems* (2020).
- [36] Alvin Chin, Jennifer Keelan, George Tomlinson, Vera Pavri-Garcia, Kumanan Wilson, and Mark Chignell. 2010. Automated delineation of subgroups in web video: A medical activism case study. *Journal of Computer-Mediated Communication* 15, 3 (2010), 447–464.
- [37] Nathan K Cobb, Amanda L Graham, and David B Abrams. 2010. Social network structure of a large online community for smoking cessation. *American journal of public health* 100, 7 (2010), 1282–1289.
- [38] Marios Constantinides, Jonas Busk, Aleksandar Matic, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E. Bardram. 2018. Personalized versus Generic Mood Prediction Models in Bipolar Disorder. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) (UbiComp ’18). Association for Computing Machinery, New York, NY,

- USA, 1700–1707. <https://doi.org/10.1145/3267305.3267536>
- [39] Seigfried-Spellar K. C. Crimmins, D. M. 2014. Peer attachment, sexual experiences, and risky online behaviors as predictors of sexting behaviors among undergraduate students. *Computers in Human Behavior* 32 (2014), 268–275. <https://doi.org/10.1016/j.chb.2013.12.012>
 - [40] Christophe Croux and Catherine Dehon. 2010. Influence functions of the Spearman and Kendall correlation measures. *Statistical methods & applications* 19, 4 (2010), 497–515.
 - [41] Maral Dadvar and Franciska De Jong. 2012. Cyberbullying detection: a step toward a safer internet yard. In *Proceedings of the 21st International Conference on World Wide Web*. 121–126.
 - [42] Gamon M.-Counts S. Horvitz E. De Choudhury, M. 2013. Predicting depression via social media. *Proceedings of the International AAAI Conference on Web and Social Media* 7, 1 (2013).
 - [43] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3267–3276.
 - [44] Munmun De Choudhury, Sanket S. Sharma, Tomaz Logar, Wouter Eekhout, and René Clausen Nielsen. 2017. Gender and Cross-Cultural Differences in Social Media Disclosures of Mental Illness. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 353–369. <https://doi.org/10.1145/2998181.2998220>
 - [45] Munmun De Choudhury, Hari Sundaram, Ajita John, Doree Duncan Seligmann, and Aisling Kelliher. 2010. "Birds of a Feather": Does User Homophily Impact Information Diffusion in Social Media? *arXiv preprint arXiv:1006.1702* (2010).
 - [46] Peter-J. de Graaf H. Nikken P. de Vries, D. A. 2016. Adolescents' social network site use, peer appearance-related feedback, and body dissatisfaction: Testing a mediation model. *Journal of youth and adolescence* 45, 1 (2016), 211–224. <https://doi.org/10.1007/s10964-015-0266-4>
 - [47] Ljiljana Despalatović, Tanja Vojković, and Damir Vukicević. 2014. Community structure in networks: Girvan-Newman algorithm improvement. In *2014 37th international convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, 997–1002.
 - [48] Lavelle M.-Pickles K. Kalorkoti C. Jaques J. Pappa S. Aylin P. Dewa, L. H. 2019. Young adults' perceptions of using wearables, social media and other technologies to detect worsening mental health: A qualitative study. *PLoS one* 14, 9 (2019), e0222655. <https://doi.org/10.1371/journal.pone.0222655>
 - [49] Andrew Disney. 2020. Social network analysis: Centrality measures. <https://cambridge-intelligence.com/keylines-faqs-social-network-analysis/>
 - [50] Nemanja Djuric, Jing Zhou, Robin Morris, Mihajlo Grbovic, Vladan Radosavljevic, and Narayan Bhamidipati. 2015. Hate speech detection with comment embeddings. In *Proceedings of the 24th international conference on world wide web*. 29–30.
 - [51] Gavin Doherty, David Coyle, and John Sharry. 2012. Engagement with Online Mental Health Interventions: An Exploratory Clinical Study of a Treatment for Depression. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 1421–1430. <https://doi.org/10.1145/2207676.2208602>
 - [52] D. M. Douglas. 2016. Doxing: a conceptual analysis. *Ethics and information technology* 18, 3 (2016), 199–210. <https://doi.org/10.1007/s10676-016-9406-0>
 - [53] Deloitte Access Economics. 2020, June. The Social and Economic Cost of Eating Disorders in the United States of America: A Report for the Strategic Training Initiative for the Prevention of Eating Disorders and the Academy for Eating Disorders. (2020, June). <https://www.hsph.harvard.edu/striped/report-economic-costs-of-eating-disorders/>
 - [54] Elizabeth V. Eikey and Madhu C. Reddy. 2017. "It's Definitely Been a Journey": A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 642–654. <https://doi.org/10.1145/3025453.3025591>
 - [55] Kristina Eriksson-Backa, Kim Holmberg, and Stefan Ek. 2016. Communicating diabetes and diets on Twitter—a semantic content analysis. *International Journal of Networking and Virtual Organisations* 16, 1 (2016), 8–24.
 - [56] Sindhu Kiranmai Ernala, Kathan H. Kashiparekh, Amir Bolous, Asra Ali, John M. Kane, Michael L. Birnbaum, and Munmun De Choudhury. 2021. A Social Media Study on Mental Health Status Transitions Surrounding Psychiatric Hospitalizations. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 155 (April 2021), 32 pages. <https://doi.org/10.1145/3449229>
 - [57] Sindhu Kiranmai Ernala, Asra F. Rizvi, Michael L. Birnbaum, John M. Kane, and Munmun De Choudhury. 2017. Linguistic Markers Indicating Therapeutic Outcomes of Social Media Disclosures of Schizophrenia. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 43 (Dec. 2017), 27 pages. <https://doi.org/10.1145/3134678>

- [58] Youssa Fettach and Lamia Benhiba. 2019. Pro-Eating Disorders and Pro-Recovery Communities on Reddit: Text and Network Comparative Analyses. In *Proceedings of the 21st International Conference on Information Integration and Web-based Application & Services (iiWAS2019)*. ACM, New York, NY, USA, 277–286. <https://doi.org/10.1145/3366030.3366058>
- [59] Jessica L. Feuston and Anne Marie Piper. 2018. Beyond the Coded Gaze: Analyzing Expression of Mental Health and Illness on Instagram. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 51 (Nov. 2018), 21 pages. <https://doi.org/10.1145/3274320>
- [60] Jannath Ghaznavi and Laramie D. Taylor. 2015. Bones, body parts, and sex appeal: An analysis of thinspiration images on popular social media. *Body Image* 14 (2015), 54–61. <https://doi.org/10.1016/j.bodyim.2015.03.006>
- [61] Paul Gilbert and Chris Irons. 2005. Focused therapies and compassionate mind training for shame and self-attacking. (2005).
- [62] Erving Goffman. 1978. *The presentation of self in everyday life*. Harmondsworth, London.
- [63] Cannon B.-Burton S. Giraud-Carrier C. Hanson, C. L. 2013. An exploration of social circles and prescription drug abuse through Twitter. *Journal of medical Internet research* 15, 9 (2013), e189. <https://doi.org/10.2196/jmir.2741>
- [64] Kelley Harper, Steffanie Sperry, and J Kevin Thompson. 2008. Viewership of pro-eating disorder websites: Association with body image and eating disturbances. *International Journal of Eating Disorders* 41, 1 (2008), 92–95.
- [65] Jenine K Harris, Alexis Duncan, Vera Men, Nora Shevick, Melissa J Krauss, and Patricia A Cavazos-Rehg. 2018. Peer reviewed: Messengers and messages for tweets that used# thinspo and# fitspo hashtags in 2016. *Preventing chronic disease* 15 (2018).
- [66] Ellison N. B. Gibbs J. L. Heino, R. D. 2010. Relationshopping: Investigating the market metaphor in online dating. *Journal of Social and Personal relationships* 27, 4 (2010), 427–447. <https://doi.org/10.1177/0265407510361614>
- [67] Chou C. Huang, Y. Y. 2010. An analysis of multiple factors of cyberbullying among junior high school students in Taiwan. *Computers in Human Behavior* 26, 6 (2010), 1581–1590. <https://doi.org/10.1016/j.chb.2010.06.005>
- [68] Smith A. R Hummel, A. C. 2015. Ask and you shall receive: Desire and receipt of feedback via Facebook predicts disordered eating concerns. *International Journal of Eating Disorders* 48, 4 (2015), 436–442. <https://doi.org/10.1002/eat.22336>
- [69] Shagun Jhaver, Sucheta Ghoshal, Amy Bruckman, and Eric Gilbert. 2018. Online Harassment and Content Moderation: The Case of Blocklists. *ACM Trans. Comput.-Hum. Interact.* 25, 2, Article 12 (March 2018), 33 pages. <https://doi.org/10.1145/3185593>
- [70] Shaohai Jiang. 2019. Functional interactivity in social media: an examination of Chinese health care organizations' microblog profiles. *Health promotion international* 34, 1 (2019), 38–46.
- [71] Karen Sparck Jones and Peter Willett. 1997. *Readings in information retrieval*. Morgan Kaufmann.
- [72] Elihu Katz and Paul F Lazarsfeld. 2017. *Personal influence: The part played by people in the flow of mass communications*. Routledge.
- [73] Yong-Hyuk Kim, Sehoon Seo, Yong-Ho Ha, Seongwon Lim, and Yourim Yoon. 2013. Two applications of clustering techniques to twitter: Community detection and issue extraction. *Discrete dynamics in nature and society* 2013 (2013).
- [74] Kara L Kimevski. 2017. *Social Media and Its Connection to the Development of Eating Disorders*. Ph.D. Dissertation. Kent State University.
- [75] Kaylee Payne Kruzan, Janis Whitlock, Natalya N. Bazarova, Katherine D. Miller, Julia Chapman, and Andrea Stevenson Won. 2020. Supporting Self-Injury Recovery: The Potential for Virtual Reality Intervention. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376396>
- [76] Emily G. Lattie, Rachel Kornfield, Kathryn E. Ringland, Renwen Zhang, Nathan Winquist, and Madhu Reddy. 2020. Designing Mental Health Technologies That Support the Social Ecosystem of College Students. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3313831.3376362>
- [77] Etienne Lave, Jean; Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- [78] Amanda Lenhart, Michele Ybarra, Kathryn Zickuhr, and Myeshia Price-Feeney. 2016. *Online Harassment, Digital Abuse, and Cyberstalking in America*. <https://datasociety.net/output/online-harassment-digital-abuse-cyberstalking/>
- [79] Kristina Lerman and Rumi Ghosh. 2010. Information contagion: An empirical study of the spread of news on digg and twitter social networks. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 4.
- [80] Stephen P Lewis and Alexis E Arbuthnott. 2012. Searching for thinspiration: the nature of internet searches for pro-eating disorder websites. *Cyberpsychology, Behavior, and Social Networking* 15, 4 (2012), 200–204.
- [81] Bussey K. Mond J. Brown O. Griffiths S. Murray S. B. Mitchison D. Lonergan, A. R. 2019. Me, my selfie, and I: The relationship between editing and posting selfies and body dissatisfaction in men and women. *Body Image* 28 (2019), 39–43. <https://doi.org/10.1016/j.bodyim.2018.12.001>
- [82] Paul Benjamin Lowry, Jun Zhang, Chuang Wang, and Mikko Siponen. 2016. Why do adults engage in cyberbullying on social media? An integration of online disinhibition and deindividuation effects with the social structure and

- social learning model. *Information Systems Research* 27, 4 (2016), 962–986.
- [83] Douglas A Luke and Jenine K Harris. 2007. Network analysis in public health: history, methods, and applications. *Annu. Rev. Public Health* 28 (2007), 69–93.
 - [84] Theo Lynn, Pierangelo Rosati, Guto Leoni Santos, and Patricia Takako Endo. 2020. Sorting the Healthy Diet Signal from the Social Media Expert Noise: Preliminary Evidence from the Healthy Diet Discourse on Twitter. *International Journal of Environmental Research and Public Health* 17, 22 (2020), 8557.
 - [85] Alexandra Marin and Barry Wellman. 2011. Social network analysis: An introduction. *The SAGE handbook of social network analysis* 11 (2011).
 - [86] Boyd D. Marwick, A. E. 2014. Networked privacy: How teenagers negotiate context in social media. *New media society* 16, 7 (2014), 1051–1067. <https://doi.org/10.1177/1461444814543995>
 - [87] Bridget Christine McHugh, Pamela J. Wisniewski, Mary Beth Rosson, Heng Xu, and John M. Carroll. 2017. Most Teens Bounce Back: Using Diary Methods to Examine How Quickly Teens Recover from Episodic Online Risk Exposure. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 76 (Dec. 2017), 19 pages. <https://doi.org/10.1145/3134711>
 - [88] Natarajan Meghanathan. 2015. Correlation coefficient analysis of centrality metrics for complex network graphs. In *Computer Science On-line Conference*. Springer, 11–20.
 - [89] Darién Miranda, Marco Calderón, and Jesus Favela. 2014. Anxiety Detection Using Wearable Monitoring. In *Proceedings of the 5th Mexican Conference on Human-Computer Interaction* (Oaxaca, Mexico, Mexico) (*MexIHC '14*). Association for Computing Machinery, New York, NY, USA, 34–41. <https://doi.org/10.1145/2676690.2676694>
 - [90] Feldhege J. Wolf M. Bauer S. Moessner, M. 2018. Analyzing big data in social media: Text and network analyses of an eating disorder forum. *International Journal of Eating Disorders* 51, 7 (2018), 656–667.
 - [91] Seunghyeon Moon, Jae-Gil Lee, and Minseo Kang. 2014. Scalable community detection from networks by computing edge betweenness on mapreduce. In *2014 International Conference on Big Data and Smart Computing (BIGCOMP)*. IEEE, 145–148.
 - [92] Hepworth J. Mulveen, R. 2006. An interpretative phenomenological analysis of participation in a pro-anorexia internet site and its relationship with disordered eating. *Journal of health psychology* 11, 2 (2006), 283–296. <https://doi.org/10.1177/13591053060061187>
 - [93] Sahiti Myneni, Kayo Fujimoto, and Trevor Cohen. 2017. Leveraging social media for health promotion and behavior change: methods of analysis and opportunities for intervention. In *Cognitive Informatics in Health and Biomedicine*. Springer, 315–345.
 - [94] Alan Neves, Ramon Vieira, Fernando Mourao, and Leonardo Rocha. 2015. Quantifying complementarity among strategies for influencers' detection on Twitter. *Procedia Computer Science* 51 (2015), 2435–2444.
 - [95] Alicia L Nobles, Caitlin N Dreisbach, Jessica Keim-Malpass, and Laura E Barnes. 2018. "Is This an STD? Please Help!": Online Information Seeking for Sexually Transmitted Diseases on Reddit. In *Twelfth International AAAI Conference on Web and Social Media*.
 - [96] 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* (Montreal QC, Canada) (*CHI '18*). Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3173574.3173987>
 - [97] Fayika Farhat Nova, Michael Ann DeVito, Pratyasha Saha, Kazi Shohanur Rashid, Shashwata Roy Turzo, Sadia Afrin, and Shion Guha. 2021. "Facebook Promotes More Harassment" Social Media Ecosystem, Skill and Marginalized Hijra Identity in Bangladesh. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–35.
 - [98] Fayika Farhat Nova, MD Rashidujjaman Rifat, Pratyasha Saha, Syed Ishtiaque Ahmed, and Shion Guha. 2019. Online sexual harassment over anonymous social media in Bangladesh. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*. 1–12.
 - [99] He D. Jeng W. Mattern E. Bowler L. Oh, J. S. 2013. Linguistic characteristics of eating disorder questions on Yahoo! Answers—content, style, and emotion. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–10.
 - [100] Kathleen O'Leary, Arpita Bhattacharya, Sean A. Munson, Jacob O. Wobbrock, and Wanda Pratt. 2017. Design Opportunities for Mental Health Peer Support Technologies. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (*CSCW '17*). Association for Computing Machinery, New York, NY, USA, 1470–1484. <https://doi.org/10.1145/2998181.2998349>
 - [101] Günce Keziban Orman, Vincent Labatut, and Hocine Cherifi. 2012. Comparative evaluation of community detection algorithms: a topological approach. *Journal of Statistical Mechanics: Theory and Experiment* 2012, 08 (2012), P08001.
 - [102] G. Overbeke. 2008. Pro-anorexia websites: Content, impact, and explanations of popularity. *Mind Matters: The Wesleyan Journal of Psychology* 3, 1 (2008), 49–62.
 - [103] Hinduja S. Patchin, J. W. 2017. Digital self-harm among adolescents. *Journal of Adolescent Health* 61, 6 (2017), 761–766. <https://doi.org/10.1016/j.jadohealth.2017.06.012>

- [104] Jessica Pater and Elizabeth Mynatt. 2017. Defining Digital Self-Harm. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 1501–1513. <https://doi.org/10.1145/2998181.2998224>
- [105] Jessica Pater, Fayika Farhat Nova, Amanda Coupe, Lauren E. Reining, Connie Kerrigan, Tammy Toscos, and Elizabeth D Mynatt. 2021. Charting the Unknown: Challenges in the Clinical Assessment of Patients' Technology Use Related to Eating Disorders. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 548, 14 pages. <https://doi.org/10.1145/3411764.3445289>
- [106] Jessica A. Pater, Brooke Farrington, Alycia Brown, Lauren E. Reining, Tammy Toscos, and Elizabeth D. Mynatt. 2019. Exploring Indicators of Digital Self-Harm with Eating Disorder Patients: A Case Study. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 84 (Nov. 2019), 26 pages. <https://doi.org/10.1145/3359186>
- [107] Jessica A. Pater, Oliver L. Haimson, Nazanin Andalibi, and Elizabeth D. Mynatt. 2016. "Hunger Hurts but Starving Works": Characterizing the Presentation of Eating Disorders Online. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) (CSCW '16). ACM, New York, NY, USA, 1185–1200. <https://doi.org/10.1145/2818048.2820030>
- [108] Jessica A. Pater, Moon K. Kim, Elizabeth D. Mynatt, and Casey Fiesler. 2016. Characterizations of Online Harassment: Comparing Policies Across Social Media Platforms. In *Proceedings of the 19th International Conference on Supporting Group Work* (Sanibel Island, Florida, USA) (GROUP '16). Association for Computing Machinery, New York, NY, USA, 369–374. <https://doi.org/10.1145/2957276.2957297>
- [109] Jessica A. Pater, Andrew D. Miller, and Elizabeth D. Mynatt. 2015. This Digital Life: A Neighborhood-Based Study of Adolescents' Lives Online. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 2305–2314. <https://doi.org/10.1145/2702123.2702534>
- [110] Jessica A. Pater, Lauren E. Reining, Andrew D. Miller, Tammy Toscos, and Elizabeth D. Mynatt. 2019. "Notjustgirls": Exploring Male-Related Eating Disordered Content across Social Media Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300881>
- [111] Carlos A Pérez-Aldana, Allison A Lewinski, Constance M Johnson, Allison A Vorderstrasse, and Sahiti Myneni. 2021. Exchanges in a Virtual Environment for Diabetes Self-Management Education and Support: Social Network Analysis. *JMIR diabetes* 6, 1 (2021), e21611.
- [112] Lorenzoni V. Andreozzi G. Mosca M. Turchetti G. Pirri, S. 2020. Topic Modeling and User Network Analysis on Twitter during World Lupus Awareness Day. *International Journal of Environmental Research and Public Health* 17, 15 (2020), 5540. <https://doi.org/10.3390/ijerph17155440>
- [113] Ava Podrazhansky, Hao Zhang, Meng Han, and Selena He. 2020. A Chatbot-Based Mobile Application to Predict and Early-Prevent Human Mental Illness. In *Proceedings of the 2020 ACM Southeast Conference* (Tampa, FL, USA) (ACM SE '20). Association for Computing Machinery, New York, NY, USA, 311–312. <https://doi.org/10.1145/3374135.3385319>
- [114] Kyle Porter. 2018. Analyzing the DarkNetMarkets subreddit for evolutions of tools and trends using LDA topic modeling. *Digital Investigation* 26 (2018), S87–S97.
- [115] Xinyu Que, Fabio Checconi, Fabrizio Petrini, and John A Gunnels. 2015. Scalable community detection with the louvain algorithm. In *2015 IEEE International Parallel and Distributed Processing Symposium*. IEEE, 28–37.
- [116] Sara Rahiminejad, Mano R Maurya, and Shankar Subramaniam. 2019. Topological and functional comparison of community detection algorithms in biological networks. *BMC bioinformatics* 20, 1 (2019), 1–25.
- [117] Abigail Karr Remick. 2010. *The effects of impression-management motivation on eating behavior in women*. Ph.D. Dissertation.
- [118] Paul Resnick and Bob Kraut. 2011. *Evidence-based social design: Mining social sciences to build online communities*. MIT University Press.
- [119] Sebastian A Rios, Felipe Aguilera, J David Nuñez-Gonzalez, and Manuel Graña. 2019. Semantically enhanced network analysis for influencer identification in online social networks. *Neurocomputing* 326 (2019), 71–81.
- [120] Elliott P. Pattison P. Robins, G. 2001. Network models for social selection processes. *Social networks* 23, 1 (2001), 1–30. [https://doi.org/10.1016/S0378-8733\(01\)00029-6](https://doi.org/10.1016/S0378-8733(01)00029-6)
- [121] Schumacher L. M. Rosenbaum D. L. Kase C. A. Piers A. D. Lowe M. R. Forman E.M. Schaumberg, K. and M.L. Butryn. 2016. The role of negative reinforcement eating expectancies in the relation between experiential avoidance and disinhibition. *Eating behaviors* 21 (2016), 129–134. <https://doi.org/10.1016/j.eatbeh.2016.01.003>
- [122] Zwillich B. Bindl M. J. Hopp F. R. Reich S. Vorderer P. Schneider, F. M. 2017. Social media ostracism: The effects of being excluded online. *Computers in Human Behavior* 73, 2017 (2017), 385–393.
- [123] Victoria Schwanda Sosik, Xuan Zhao, and Dan Cosley. 2012. See Friendship, Sort of: How Conversation and Digital Traces Might Support Reflection on Friendships. In *Proceedings of the ACM 2012 Conference on Computer Supported*

- Cooperative Work* (Seattle, Washington, USA) (CSCW '12). Association for Computing Machinery, New York, NY, USA, 1145–1154. <https://doi.org/10.1145/2145204.2145374>
- [124] Gayen A. Ganguly N. Dandapat S. K. Chandra J. Sequeira, R. 2019. A Large-Scale Study of the Twitter Follower Network to Characterize the Spread of Prescription Drug Abuse Tweets. *IEEE Transactions on Computational Social Systems* 6, 6 (2019), 1232–1244. <https://doi.org/10.1109/TCSS.2019.2943238>
 - [125] Leslie Regan Shade. 2003. Weborexics: The ethical issues surrounding pro-ana websites. *Acm Sigcas Computers and Society* 33, 4 (2003), 2.
 - [126] Chiu C. M. Chang C. C. Shiue, Y. C. 2010. Exploring and mitigating social loafing in online communities. *Computers in Human Behavior* 26, 4 (2010), 768–777. <https://doi.org/10.1016/j.chb.2010.01.014>
 - [127] Will Simm, Maria Angela Ferrario, Adrian Gradinar, Marcia Tavares Smith, Stephen Forshaw, Ian Smith, and Jon Whittle. 2016. Anxiety and Autism: Towards Personalized Digital Health. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 1270–1281. <https://doi.org/10.1145/2858036.2858259>
 - [128] Smith P. K. Frisén A. Slonje, R. 2013. The nature of cyberbullying, and strategies for prevention. *Computers in human behavior* 29, 1 (2013), 26–32. <https://doi.org/10.1016/j.chb.2012.05.024>
 - [129] Szongott C. Henne B. Von Voigt G. Smith, M. 2012. Big data privacy issues in public social media. In *Proceedings of the 6th IEEE international conference on digital ecosystems and technologies (DEST '12)*. IEEE, New York, NY, USA, 1–6. <https://doi.org/10.1109/DEST.2012.6227909>
 - [130] Peter Snyder, Periwinkle Doerfler, Chris Kanich, and Damon McCoy. 2017. Fifteen Minutes of Unwanted Fame: Detecting and Characterizing Doxing. In *Proceedings of the 2017 Internet Measurement Conference* (London, United Kingdom) (IMC '17). Association for Computing Machinery, New York, NY, USA, 432–444. <https://doi.org/10.1145/3131365.3131385>
 - [131] Shaina J Sowles, Monique McLeary, Allison Optican, Elizabeth Cahn, Melissa J Krauss, Ellen E Fitzsimmons-Craft, Denise E Willfey, and Patricia A Cavazos-Rehg. 2018. A content analysis of an online pro-eating disorder community on Reddit. *Body image* 24 (2018), 137–144.
 - [132] Samuel Alan Stewart and Syed Sibte Raza Abidi. 2012. Applying social network analysis to understand the knowledge sharing behaviour of practitioners in a clinical online discussion forum. *Journal of medical Internet research* 14, 6 (2012), e170.
 - [133] Pranav Suri and Nihar Ranjan Roy. 2017. Comparison between LDA & NMF for event-detection from large text stream data. In *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICIT)*. IEEE, 1–5.
 - [134] Didi Surian, Dat Quoc Nguyen, Georgina Kennedy, Mark Johnson, Enrico Coiera, and Adam G Dunn. 2016. Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. *Journal of medical Internet research* 18, 8 (2016), e232.
 - [135] G Suryateja and P Saravanan. 2018. Optimization of Community Detection in Twitter Social Network. *International Journal of Pure and Applied Mathematics* 118, 20 (2018), 745–758.
 - [136] Fernandez-Luque L. Jian W. S. Li Y. C. Crain S. Hsu M. H. ... Liou D. M. Syed-Abdul, S. 2013. Misleading health-related information promoted through video-based social media: anorexia on YouTube. *Journal of medical Internet research* 15, 2 (2013), e30. <https://doi.org/10.2196/jmir.2237>
 - [137] Lauren C. Taylor, Kelsie Belan, Munmun De Choudhury, and Eric P. S. Baumer. 2021. Misfires, Missed Data, Misaligned Treatment: Disconnects in Collaborative Treatment of Eating Disorders. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 31 (April 2021), 28 pages. <https://doi.org/10.1145/3449105>
 - [138] Marika Tiggemann, Owen Churches, Lewis Mitchell, and Zoe Brown. 2018. Tweeting weight loss: A comparison of #thinspiration and #fitspiration communities on Twitter. *Body Image* 25 (2018), 133–138.
 - [139] Nicholas J.-Larsen M. E. Firth J. Christensen H. Torous, J. 2018. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evidence-based mental health* 21, 3 (2018), 116–119. <https://doi.org/10.1136/eb-2018-102891>
 - [140] S. Vallor. 2012. Flourishing on facebook: virtue friendship new social media. *Ethics and Information technology* 14, 3 (2012), 185–199. <https://doi.org/10.1007/s10676-010-9262-2>
 - [141] Wright-M. F. Wachs, S. 2019. The moderation of online disinhibition and sex on the relationship between online hate victimization and perpetration. *Cyberpsychology, Behavior, and Social Networking* 22, 5 (2019), 300–306. <https://doi.org/10.1089/cyber.2018.0551>
 - [142] Rui Wang, Weichen Wang, Min Hane Aung, Dror Ben-Zeev, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Emily A. Scherer, and Megan Walsh. 2018. Predicting Symptom Trajectories of Schizophrenia Using Mobile Sensing. *GetMobile: Mobile Comp. and Comm.* 22, 2 (Sept. 2018), 32–37. <https://doi.org/10.1145/3276145.3276157>

- [143] 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* (Cambridge, United Kingdom) (*WSDM '17*). Association for Computing Machinery, New York, NY, USA, 91–100. <https://doi.org/10.1145/3018661.3018706>
- [144] Tao Wang, Markus Brede, Antonella Ianni, and Emmanouil Mentzakis. 2019. Characterizing dynamic communication in online eating disorder communities: a multiplex network approach. *Applied Network Science* 4, 1 (2019), 1–22.
- [145] Yuda Wang and Gang Li. 2018. The spreading of information in online social networks through cellular automata. *Complexity* 2018 (2018).
- [146] Massagli M.-Frost J. Brownstein C. Okun S. Vaughan T. Bradley R. Wicks, P. and J. Heywood. 2010. Sharing health data for better outcomes on PatientsLikeMe. *Journal of medical Internet research* 12, 2 (2010), e19. <https://doi.org/10.2196/jmir.1549>
- [147] Pamela Wisniewski, Haiyan Jia, Na Wang, Saijing Zheng, Heng Xu, Mary Beth Rosson, and John M. Carroll. 2015. Resilience Mitigates the Negative Effects of Adolescent Internet Addiction and Online Risk Exposure. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 4029–4038. <https://doi.org/10.1145/2702123.2702240>
- [148] Pamela Wisniewski, Heng Xu, Mary Beth Rosson, Daniel F. Perkins, and John M. Carroll. 2016. Dear Diary: Teens Reflect on Their Weekly Online Risk Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 3919–3930. <https://doi.org/10.1145/2858036.2858317>
- [149] Janis Wolak, Kimberly J Mitchell, and David Finkelhor. 2007. Does online harassment constitute bullying? An exploration of online harassment by known peers and online-only contacts. *Journal of adolescent health* 41, 6 (2007), S51–S58.
- [150] Qin Wu, Xingqin Qi, Eddie Fuller, and Cun-Quan Zhang. 2013. “follow the leader”: A centrality guided clustering and its application to social network analysis. *The Scientific World Journal* 2013 (2013).
- [151] Weiai Wayne Xu, I-Hsuan Chiu, Yixin Chen, and Tanuka Mukherjee. 2015. Twitter hashtags for health: applying network and content analyses to understand the health knowledge sharing in a Twitter-based community of practice. *Quality & Quantity* 49, 4 (2015), 1361–1380.
- [152] Dawei Yin, Zhenzhen Xue, Liangjie Hong, Brian D Davison, April Kontostathis, and Lynne Edwards. 2009. Detection of harassment on web 2.0. *Proceedings of the Content Analysis in the WEB 2* (2009), 1–7.
- [153] Adam G Zimmerman and Gabriel J Ybarra. 2016. Online aggression: The influences of anonymity and social modeling. *Psychology of Popular Media Culture* 5, 2 (2016), 181.

A APPENDIX

A.1 LDA Table

A.2 NMF Table

Received July 2021; revised September 2021; accepted October 2021

Primary Theme	Secondary Theme(s)	Representative Keywords
Community Build/ Maintenance	-Motivational conversation -Coping mechanism	edtwrt, retweet, motivation, friend, anorexia
Best Practices and plans	-Sharing best practices (e.g. melatonin to stay asleep so can't eat; laxatives for IBS; fasting tips) -Daily updates/Plan	group, feel, tips, binge, today
Specific ED Inspiration	-Desired ideals/Longing for desired actions/Celebrity idolization - Malespo/gender specific inspiration	body, goal, malespo, wanna, wait
Weight loss setting/status	-Accountability/status -Weight loss	weight, lose, pound, goal, loss
Thought Processes during active ED	-Antagonistic rhetoric/negative thought process (Struggle, Meanspo) -Self-realization/feelings	Fuck, hungry, look, disgust, feel
Presentation of (ED) self	-Outwards exhibitions of diet, meal-prep and status -Negative reflections of others to build up oneself -Gender/ethnicity (internal and external)	eat, starve, hate, self, people
Fitness/Diet (Diary type discussions related to process of fitness/diet; more intrinsic)	-Fitness Regime -Seeking Advice/Q/A -Resource sharing (including specific social media content providers)	calorie, water, workout, fast, know
Personal reflection on body image	-Tone (Neutral, sad, frustrated, defeated) -Specific body traits	skinni, want, imagine, thigh, leg
Engagement with sub-communities	-Reaching sub-communities for ED (international and known sub-communities) - ED journey (withdraws/relapse) - Impacts with community (toxicity/disorder)	bslyw, abwtb, waniliowemleko, disorder, toxic
ED experiences and values	-Mis-category of ED hashtags -Ed related memes -Valuing ED lifestyle (Earning/deserving)	edtwrt, edmem, deserve, hard, food

Table 10. Primary and Secondary Themes from LDA

Topic	Primary Theme	Secondary Theme(s)	Keywords
0	Community based characteristics and curating	- Community gate-keeping/ building between Twitter to Tiktok -General thinspo	edtwit, mutuals, dailythin, calorie, workout
1	Popularity of major community member	-User has been suspended and content analysis can't be conducted	TikTokDiets, block, instagram, leave
2	Disorder specific community engagement	-Proana Thinspo -Proana diet/fasting (calorie counting/exercise)	proana, promia, anorexia, edtwitt, diet
3	Community building	-Membership status (active/inactive) -Expanding the community (new-connections/new-additions)	like, mutual, follow, friend, active
4	ED related Aspiration	-Body traits (including Malespo/ poc inspiration/clothing) -Diet/fast updates -Personal reflections/accountability	skinni, thinspir, edtwitt, girl, leg
5	Weight Loss	-Community support/tips (motivations) -Updates (struggle/status/plan)	lose, weight, pound, goal, know
6	Support/Motivations during active ED	-Criticism/Meanspo (inward/outward) -Motivation/support (inward/outward) -Seeking meanspo	meanspo, send, edtwitt, motivation, need
7	Targeted engagements and presentations of ED journey	-Specific components of ED (Body traits/clothing/disciplinary motivations) -Group chat (social media/ fasting groups)	group, friend, toxic, fast, chat
8	ED specific reflections	-Daily ED reflections (emotions/events) -Comparison reflecting ED self (outward-celebrity/personal relationships) -ED memes -(Positive) Encouragement/tips (outwards)	think, edmem, funnier, everytim, feel
9	Desired ED lifestyle	-Tone (frustration/self-hate) -Extreme idealization of ED values (extreme: hard work/goal) -Celebrity Idolization/goal	body, fuck, hate, kill, wish

Table 11. NMF Primary and Secondary Topics