Supporters First: Understanding Online Social Support on Mental Health from a Supporter Perspective

MEEYUN KIM, Incheon National University, Republic of Korea
KOUSTUV SAHA, Microsoft Research, Canada
MUNMUN DE CHOUDHURY, Georgia Institute of Technology, U.S.A
DAEJIN CHOI*, Incheon National University, Republic of Korea

Social support or peer support in mental health has successfully settled down in online spaces by reducing the potential risk of critical mental illness (e.g., suicidal thoughts) of support-seekers. While the prior work has mostly focused on support-seekers, particularly investigating their behavioral characteristics and the effects of online social supports to support-seekers, this paper seeks to understand online social support from supporters' perspectives, who have informational or emotional resources that may affect support-seekers either positively or negatively. To this end, we collect and analyze a large-scale of dataset consisting of the supporting comments and their target posts from 55 mental health communities in Reddit. We also develop a deep-learning-based model that scores informational and emotional support to the supporting comments. Based on the collected and scored dataset, we measure the characteristics of the supporters from the behavioral and content perspectives, which reveals that the supporters tend to give emotional support than informational support and the atmosphere of social support communities tend also to be emotional. We also understand the relations between the supporters and the support-seekers by giving a notion of "social supporting network", whose nodes and edges are the sets of the users and the supporting comments. Our analysis on top users by out-degrees and in-degrees in social supporting network demonstrates that heavily-supportive users are more likely to give informational support with diverse content while the users who attract much support exhibit continuous support-seeking behaviors by uploading multiple posts with similar content. Lastly, we identified structural communities in social supporting network to explore whether and how the supporters and the support-seeking users are grouped. By conducting topic analysis on both the support-seeking posts and the supporting comments of individual communities, we revealed that small communities deal with a specific topic such as hair-pulling disorder. We believe that the methodologies, dataset, and findings can not only expose more research questions on online social supports in mental health, but also provide insight on improving social support in online platforms.

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

Additional Key Words and Phrases: peer support, social support, mental health, online communities, Reddit

*Corresponding author

Authors' addresses: Meeyun Kim, tr0up2r@inu.ac.kr, Incheon National University, Incheon, Republic of Korea; Koustuv Saha, koustuvsaha@microsoft.com, Microsoft Research, Montréal, Québec, Canada; Munmun De Choudhury, munmund@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, U.S.A; Daejin Choi, djchoi@inu.ac.kr, Incheon National University, Incheon, Republic of Korea.

2573-0142/2023/4-ART92 $15.00
https://doi.org/10.1145/3579525

ACM Reference Format:
Meeyun Kim, Koustuv Saha, Munmun De Choudhury, and Daejin Choi. 2023. Supporters First: Understanding Online Social Support on Mental Health from a Supporter Perspective. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 92 (April 2023), 28 pages. https://doi.org/10.1145/3579525

1 INTRODUCTION

Social support or peer support in mental health, providing psychological or informational resources to individuals with mental illnesses such as depression, eating disorder, or even suicide ideation [57], has increasingly appeared in online spaces like Reddit, 7Cups, or Talklife [46, 64]. In these platforms, an individual who requests help, often called a support-seeker, indicates their mental health status, suffering situation, or current concerns affecting mental illness. Supporters then give the support-seeker informational or emotional support including informational materials, advice, empathy, or acknowledgement [56]. With the anonymity property that encourages individuals with mental illness to less hesitate to reveal their mental health status or self-disclosure [2, 3, 18, 21], and the increasing quality of informational or emotional support, social support in online spaces, dubbed as online social support, has been considered as an effective way to reduce the potential risk of critical mental illness (e.g., suicidal thoughts) of support-seekers [19, 41].

The great upsurge of online social support has spurred the research communities to explore the characteristics of online social support in diverse topics, including HIV [32], cancer [58, 67], or sexual experience/abuse [2, 26, 48]. In the mental health domain, there has been much effort to measure online social support, particularly support-seeking behaviors [56], psychological and behavioral changes of support-seekers [1, 12, 19, 43, 46, 66], or communication patterns among users [55]. Concentrating on support-seekers, these studies have provided valuable insights into understanding what supports are required, how effective, and what factors (e.g., linguistic accommodation of support-seeking posts) are related to attracting better support.

Although prior research has leveraged constructive ways to care for individuals' mental health problems, relatively little work has focused on "supporting users or supporters". A few recent studies have identified diverse roles such as informational/emotional support providers (i.e., supporters) based on statistical methods [39, 67], but the behavioral investigation for supporters has still remained. Since supporting users have potential resources that may affect support-seekers either positively or negatively [39], we argue that understanding the characteristics of supporters can be an important study to understand how to provide better support to individuals with mental illness, which in turn can result in improvement of online social support system.

This paper seeks to understand online social support in mental health from the supporter's perspective. To the best of our knowledge, it is the first attempt to quantitatively analyze the characteristics of online social support from a supporter's perspective in online communities for mental health. In particular, we summarize the research questions as follows:

- **RQ1 (General Behaviors):** How do supporters help support-seekers? How diverse topics are supporters engaged with and how similar are their responses?
- **RQ2 (Active Users):** What are the significant characteristics of heavily-supportive users (i.e., supporters who helps a number of support-seekers) and attractive support-seekers who elicit much support?
- **RQ3 (User Groups):** Whether and how supporters and support-seekers are grouped? Is there a specific topic mentioned in a user group? If yes, what topics and supporting resources are dealt within the group?

To answer these questions, we collect and analyze a large scale of the supporting comments and their target posts from 55 mental health communities on Reddit, a popular online community where support-seekers upload posts with their mental health status and illness, and the supporters help by giving comments. Using the ground-truth dataset provided by [56], which includes the comments with the scores for informational support (IS) and emotional support (ES), we train a deep learning-based model to assign the IS and ES scores to each supporting comment in the collected dataset. We then measure the characteristics of the supporters from the behavioral and content perspectives to address RQ1. To answer RQ2 and RQ3, we model the interactions between supporters and support-seekers as "social supporting networks", whose nodes and edges are the sets of the users and the supporting comments, respectively. On the social supporting network, we identify heavily-supportive users and the attractive support-seekers who receive much support and characterize the users from both the behavioral and content perspectives, which addresses RQ2. By computing the structural communities (i.e., user groups) in social supporting network and conducting LDA analysis for individual communities, we investigate whether and how users can be grouped in terms of topical characteristics, which is an answer for RQ3.

We believe that the analysis results in this paper can provide valuable insights into improving online social support in mental health. For example, a key finding of this paper is that supporters tend more to give emotional support than informational support. Based on the finding, not only support-seekers can use online social media as a space to receive emotional support, but also clinicians can encourage individuals to leverage social media for providing emotional support. We also find that informational support given by heavily-supportive users is important and effectively considered, and the user groups dealing with specific topics related to mental health (e.g., Trichotillomania) can be identified, which can be used to design supporter recommendation systems for support-seekers. We discuss the implications of this work.

2 RELATED WORK

2.1 Social Support: Definition, Types, Effect, and Roles

Since the 1970s, social support has steadily been paid attention to by research communities. A notable definition of social support was suggested by Kaplan et al. [31], who described social support as "metness or gratification of a person's basic social needs including approval, esteem, and succorance by significant others". With the definition, there have been a few studies to explore the properties or types of social support. Caplan [7] detached emotional and instrument support from others in his work[7, 8]. Similarly, Kahn and Antonucci [30] defined three classes of social support: (i) affect that involves the expression of caring and emotional intimacy, (ii) affirmation that provides information about the rightness of one's actions or thoughts, and (iii) aid showing the availability of direct help. A popular classification of social support was proposed by Goldsmith et al. [23], who suggested five categories of social support: informational, emotional, tangible aid, social network, and esteem. The classification method has widely been employed in many prior studies, mainly focusing on informational and emotional support [19, 53, 56, 62]. Our work is within the context of these studies – the same classification scheme and quantification method are used in this work. Although the classification method above is popularly used, we note that the different categorization methods were used in a part of online social support research regarding its purpose [33, 45].

Another large body of work has contributed to understanding roles and effects of social support [15, 37]. One of the well-known effects of social support is "buffering hyphothesis" [14, 15], which assumes that social support protects people from the pathogenic effects of a stressful life event. Note that the buffering effect was verified and complemented by a large number of prior studies [4, 36, 47, 57, 61, 63]. Dalgard et al. [16] confirmed the existence of buffering effect by social support and showed that the effect becomes stronger for depression. On the other hand, Thoits [60] complemented the buffering hypothesis by emphasizing the importance of clear conceptualization on the relations among psychological disturbance, social support, and life events (causing stress). Focusing only on the social supports perceived by a recipient, Lakey and Orehek [36] proposes Relational regulation theory (RRT), which claims that relations between perceived (social) support and its effect on mental health primarily emerge from affectively consequential conversations about people's daily lives, not about how to cope with stress.

In addition to the theoretical contributions, there also have been a few empirical studies to explore what social support affects in practice [17, 22, 28]. For example, George et al. [22] investigated what and how social support are associated with the outcomes of depression by analyzing in-patients data, revealing that subjective social support was most strongly associated with recovery of depression [22]. Conducting a web-based survey on college students, Hefner and Eisenberg [28] found that the students given low-quality social support were more likely to experience mental health problems, such as an increased risk of depression. On the other hand, Davis et al. [17] investigated the relations between supportive behavior and adjustment following rape by analyzing data from 105 victims of rape, attempted rape, or aggravated sexual assault, which demonstrated that high levels of supportive behavior are not guaranteed the better adjustment to rape. A few studies have contributed to revealing whether social support given by specific social ties (i.e., family members or friends) is effective [38, 54, 59]. Stice et al. [59] showed that parental support in early adolescence reduces the risk of depression while Leung et al. [38] found that emotional involvement (i.e., support) by family members was inversely correlated to the degree of depression and anxiety of elderly people [38]. In [54], the authors interviewed 16 survivors of Columbine High School Massacre that happened at April 20, 1999, and found that the social support from a survivor network consisting of people with similar experience is most effective, while the well-intended support from outside of the network is relatively unhelpful. Caron et al. [9] used a sequential regression method to find the predictors of distress and found that emotional support and the presence of people perceived as stressful together are the key predictors of individuals' distress. While these studies have provided valuable insights on understanding social support appeared in offline spaces, the results of the studies can hardly be generalized to online social support due to the different nature of social supports between online and offline spaces. Contrary to offline spaces, social support in online platforms can be given by unspecified supporters rather than acquaintances, which requires a different approach for analysis. In this paper, we selected an online space, Reddit, to understand online social support from supporters' perspectives.

2.2 Online Social Support for Health

With the increasing usage of online platforms, social support has been expanded to online spaces, dubbed as "Online Social Support", in which support-seekers and supporters interact in online communities. The provision of anonymity and temporal and spatial efficiency of communication has proliferated online social support, which has, in turn, attracted the research community to understand online social supports in diverse topics including cancer [58, 65, 67], HIV [32], pregnancy [25], and sexual experiences [26, 48]. Yang et al. [67] identified multiple roles of peers such as emotional support provider, welcomer, or story sharer based on the behavioral characteristics of users in Cancer Survivor Network, online social support platforms for cancer, and found that a large portion of peers is either emotional support provider or welcomer. On the other hand, Levonian et al. [39] analyzed the dataset from CaringBridge, an online service to provide social support for cancer patients and caregivers who have a responsibility to care for the patients in an offline space. On three different roles of users, caregiver, patient, and mixed, the authors investigated the characteristics of interaction and relations among the different roles of peers, which showed that the users of the same role are more likely to form interactive relationships. In further work on the same dataset, Smith et al. [58] defined friends, family, and acquaintances (FFA) as another role. The authors then explored what social support is importantly and perceived by different roles of users, which found that informational support is perceived important by patients and caregivers rate while FFA think emotional support is more important. The authors in [65] explored the effect of Bosom Buddies, a group of online social support for breast cancer, and demonstrated that the supporting group effectively reduces participants' depression, perceived stress, and cancer-related trauma. Gui et al. [25] measured the topics of support-seeking posts in BabyCenter, a social supporting service for pregnancy. On the other hand, Karusala et al. [32] investigated the interaction patterns of online social support for HIV through a WhatsApp dataset. Online social support for sexual experience has also been studied in prior work [26, 48]. In this paper, we target online social support for mental health.

2.3 Online Social Support for Mental Health

In recent years, there has been much investigation on online social support for mental health, particularly the characteristics of support-seekers [3, 35] or their posts [33, 34]. Andalibi et al. [3] characterized the support-seeking behaviors of the users who suffered from sexual abuse history on Reddit and found that greater anonymity drives more engagement of the users in seeking support. Kushner and Sharma [35] analyzed temporal patterns of supporting activities in TalkLife, revealing that activity intensity of online social support is different over time. Kruzan et al. [33] defined three types of support-seeking posts of self-injuring users and measured what supports are differently given to each type of the posts.

Another rich branch of the work has paid attention to the effects of online social supports from psychological or behavioral perspectives [12, 19, 50, 55, 69]. De Choudhury and Kiciman [19] analyzed the language of the social support in mental health subreddits on Reddit and found the treatment keywords related to reducing the risk of individuals' suicide ideation. Pruksachatkun et al. [46] compared the posts before/after providing social support to observe the cognitive changes of support-seekers. Investigating online peer supports given to people living with cancer, Allison et al. [1] showed that their mental health status can be improved by reducing the risk of mental illness like depression. Sharma et al. [55] investigated the characteristics of the seeking and supporting activities on both Reddit and TalkLife datasets and analyzed the relations between these factors and the staying probability of support-seekers in the engaged online communities while Yang et al. [66] related the staying length and existence/types of received supports to users' commitment to communities. Similarly, Chen and Xu [12] quantified the probability and the speed of support seekers' returns to communities based on the types of received support.

While the prior studies have provided great insights on understanding support-seeking behaviors and the effects of online social support, our work explores online social support from a supporter's perspective. To the best of our knowledge, it is the first attempt to quantitatively analyze the behavioral and content characteristics of supporters, who have potential resources that may affect support-seekers.

3 DATA PREPARATION

3.1 Data Collection and Preprocessing

We investigate the peer-supporting behaviors on mental health (MH) by collecting the posts and the comments uploaded in 55 MH-related subreddits, including r/abuse, r/depression, r/ptsd, and r/Anxiety, which were comprehensively investigated and listed by prior research [51, 56]. In these subreddits, the support-seekers upload a post that indicates their own status of mental illness or suffering situation. Other subreddit members (supporters) then reply to the comments, including advice that is believed to address and solve the given problems. For example, a supporter, who also suffered from a similar situation in the past, can share a coping way or empathize with encouraging the support-seeker. To retrieve the support-seeking posts and their associated supports (i.e., comments) in Reddit, we used Pushshift [1], a service that provides APIs for systematic acquisition of the posts, comments, subreddits, and users since 2008. Using the publicly available PushShift API, we collected all the posts and their associated comments uploaded to 55 subreddits in 2020. Here, we only selected the first-level comments that are direct replies to the original post. The rationale of this approach is in line with the prior work [13], which reported that the topics of the nested comments (i.e., comments of comments) tend to become more distant from the one of the original posts in threaded conversations. Note that the same approach was also taken from the prior study [56]. Lastly, we filtered out the deleted or removed posts/comments, including no content texts.

From the collected dataset, we excluded the posts and the comments written by the authors whose behaviors are not relevant to social support. In particular, we removed all the posts and their associated comments uploaded by (i) the moderating bots officially listed in subreddits, whose comments are usually the guidance or the rules about post uploads, (ii) other bots that user ids are ended with 'bot', and (iii) the deleted accounts at the data collection period. In addition, we manually marked four writers as invalid, whose posts elicited more than 13 K comments, but the content is irrelevant to mental health problems. The example titles of the posts are "Let's post good news on the coronavirus here", "Coronavirus Megathread", and "Live Chat for (06/20/20) Please read the rules in the sidebar and in the body of text below". The total numbers (portions) of the ignored users, posts, and comments are 618 (0.07%), 31,244 (2.47%), and 802,635 (27.58%), respectively.

3.2 Assigning Informational/Emotional Support Scores

We next classify the collected social supports (i.e., comments) in terms of informational support (IS) and emotional support (ES) and compute IS/ES scores for the comments. The classifying scheme was originated from the "Social Support Behavioral Code", proposed by Cutrona and Suhr [23], which defines five types of support, including informational, emotional, instrumental, esteem, and network support. IS and ES have mainly been paid attention to and employed as a lens of analysis in the prior work for online social support [19, 50, 53, 56, 62]. We adopt the classification and quantifying scheme proposed by Sharma and De Choudhury [56], which scores for IS/ES in Likert scale (1=least supportive, 3=most supportive). This proposed scheme has also been adopted in other recent studies [50, 52].

To calculate the IS/ES scores for the collected comments, we develop a deep-learning model based on Bidirectional Encoder Representations from Transformers (BERT) [20], a popular and successful model in extracting comprehensive linguistic features from a given text. Here, we use a pre-trained $BERT_{base}$, which consists of 12 layers and 12 attention heads using 110 M parameters and learns from a large amount of text collected from BookCorpus and Wikipedia. [2] Adding a fully-connected layer after $BERT_{base}$, the model labels the IS/ES score to each comment in Likert scale. Note that this approach, which *fine-tunes* a pre-trained BERT model, has been popularly taken for text classification problems (e.g., sentiment classification). The epoch counts, batch size, learning rate, and epsilon are set to 10, 3, 1e-5, and 1e-8, respectively. We used AdamW [40] to optimize the proposed model.

[1]https://github.com/pushshift/api

[2]We downloaded a pre-trained $BERT_{base}$ model from here: https://huggingface.co/bert-base-uncased

Table 1. The statistical description of collected dataset.

# posts	1,235,395
# comments	2,107,308
# users	691,426
Collection period	1 Jan. 2020 - 31 Dec. 2020
Average # words (post titles/comment texts)	8.92/53.07
Comment ratios (%) by IS scores (1/2/3)	65.7/24.0/10.3
Comment ratios (%) by ES scores (1/2/3)	36.5/38.7/24.8

The proposed model is then trained (or fine-tuned) with the dataset used in [56]. That is, the dataset consists of 400 comments, which were inspected and assigned IS/ES scores in Likert scale by two human annotators. The numbers of comments whose IS/ES scores are 1, 2, and 3 are 185/135, 131/134, and 81/128. Dividing the ground-truth dataset into training and test sets with 8:2 ratio, we separately trained two fine-tuned BERT models for IS and ES using the training set. The performance of the proposed model evaluated by the test set outperforms the proposed model in the original study; our model achieves 72.5%/71.3% and 0.726/0.715 for accuracy and F1-scores, respectively, in terms of IS/ES scores while the ones of the original model are 72%/59% and 0.68/0.57, respectively, which demonstrates the proposed model in this work can score IS/ES for the social supports accurately. Using these fine-tuned models, we finally label IS/ES scores in all the newly collected comments in this work.

3.3 Data Description

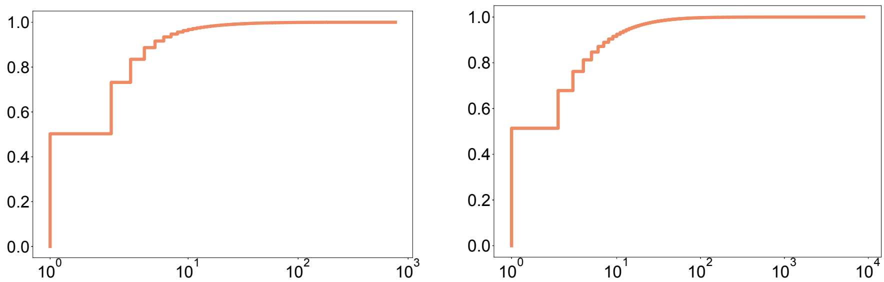
Table 1 shows the description of the dataset finally obtained. The total numbers of the posts and comments are 1,235,395 and 2,107,308, respectively, written by 691,426 anonymous users. The average numbers of words in the post titles and the comment texts are 8.92 and 53.07, respectively. The numbers and portions of the comments whose IS scores are 1, 2, and 3 are 2,924,507 (65.7%), 1,067,719 (24.0%), and 459,978 (10.3%), while the ones for ES are 1,626,356 (36.5%), 1,721,821 (38.7%), and 1,104,027 (24.8%).

Ethical Considerations: We explicitly declare the ethical consideration on the dataset and analysis. Since the collected dataset is publicly available and already de-identified, the study did not constitute human subjects research, and therefore the ethical review board of the research institute considered the research exempt.

4 GENERAL BEHAVIOR AND CONTENT ANALYSIS FOR SUPPORTERS

We start by analyzing the general characteristics of supporters from behavioral and content perspectives, which answers RQ1. In particular, we first measure the basic distributions of supporting activities, and then explore (i) how diverse topics a supporter is engaged with and (ii) how similar his/her supports are, and (iii) what atmospheres are formed in online communities for social support, which can provide insights on what social supports are given and how, in online spaces.

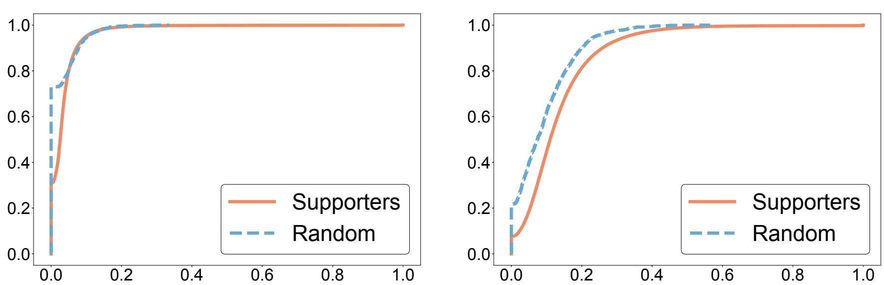
Figure 1 shows the distribution of the number of supporters in a post and the number of comments written by a supporter. Both distributions indicate a heavy tail that spans several orders of magnitudes. For example, while around half of the support-seeking posts are replied to by only a supporter, 0.1% of posts attract more than 92 supporters, which implies that the attraction is highly skewed to a few posts. The distribution of the number of the comments exhibits a larger tail; only 0.01% of the supporters uploaded more than 100 comments, while 51.36% of the users give support only once. It is also observed that the distribution of the number of the participating posts of a



(a) Number of Supporters (b) Number of Comments

Fig. 1. The distributions of the number of the supporters in a support-seeking post and the number of comments made by a supporter.

supporter is almost identical to Figure 1b, indicating that it is rare for a supporter directly replies to a post twice.



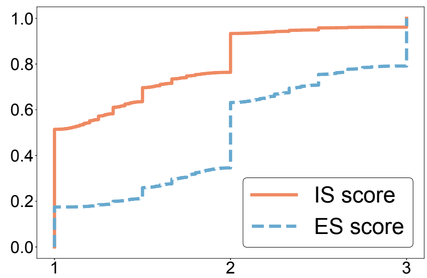
(a) Titles of Participating Posts (b) Comment Texts

Fig. 2. The distributions of the TF-IDF similarity scores among the comment texts and the titles of the participating posts.

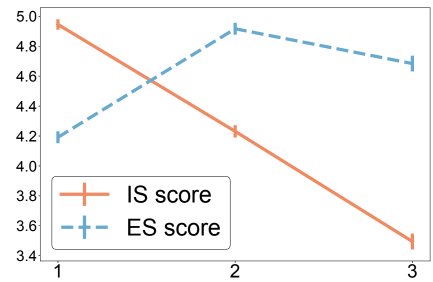
We next investigate how diverse topics a supporter is engaged with and how similar the supporter's responses are, which can indicate the characteristics of participating behaviors of supporters in terms of topics. That is, if a supporter is involved in the posts with similar topics, we can conjecture that the supporters tend to participate in the specific topics of social support. To explore this, we compute the average scores of document similarity among (i) the titles of the support-seeking posts that each supporter replied to and (ii) the comment texts made by the user, respectively. To this end, we first compute Term Frequency-Inverse Document Frequency (TF-IDF) vectors of the post titles (or comment texts) related to each supporter, whose element w_i is calculated as:

$$w_i = tf_i \cdot (\log \frac{1+n}{1+df_i} + 1) \tag{1}$$

where n, tf_i, and df_i are the number of posts (comments) that a supporter replies to (response), the term frequency, and inverse document frequency of i-th word in a document, respectively. We then calculate the average score of cosine similarity of all the pairs of the calculated TF-IDF vectors. The score is in the range of [0, 1], where the score is 1 when the post titles (comment texts) associated with a given supporter are identical and 0 when no words are overlapped. Note that we ignore the supporters engaged with less than two posts (comments).



(a) Average IS/ES scores



(b) Average voted scores

Fig. 3. The distributions of the average IS/ES scores for the supporters and the number of the received votes in terms of IS/ES scores.

Figure 2 shows the distribution of the similarity scores of the post titles and comment texts. Here, we also calculate the average scores of TF-IDF similarity for 1,000 randomly-selected pairs (of either the post titles or the comment texts) as a comparison purpose (denoted by Random). As shown in Figure 2, the average scores for both the post titles and the comment texts are higher than the ones of the random pairs (p<.001 of two-sample t-tests in both cases). The similarity scores of the post titles for 73% random pairs are zero while the same case is observed for only 30.92% supporters. In addition, the distribution of the post titles for Supporters shows longer tails, compared with the Random case. The tendency is significantly stronger in the case of comment texts. These results imply that the supporters are likely to participate in similar topics with a similar response. Note that the comment texts show a higher similarity than the post titles, which is caused by the length of the text. That is, since the comment texts are longer than the post titles, the probability of the occurrence of the duplicated words in a pair of the comment texts is higher.

We further explore the general atmosphere of online social support by analyzing what support are mainly given by supporters and encouraged by community audiences, which can provide valuable guidance that how to use online social support to support-seekers. To this end, we measure the average IS/ES scores of the comments and the voted scores given to a supporter, described in Figure 3. The voted scores are calculated as the difference between the number of upvotes and downvotes, indicating how much the provided support elicits empathy or agreement from community audiences. Note that a user upvotes when he/she/they feel sympathy/empathy or agree with the described context, while they are likely to downvote when the content does not contribute to a constructive conversation. Overall, the average ES scores are higher than IS scores; the average IS scores of the comments from more than half of supporters are one, while around 70% supporters upload the comments whose ES scores are equal to or more than two, indicating that online social supports tend more to be emotional. Note that such trend was also observed in social support for other topics such as online cancer forums [67].

Figure 3b indicates an interesting pattern between average voted scores and IS/ES scores; there is a significantly negative correlation between the average voted scores and the IS scores. In contrast, the average voted scores increase as the ES scores increase from 1 to 2, then slightly decrease. These results imply that online social support for mental health forms an emotional atmosphere – the emotional support is more likely to attract empathy from the community audiences, and the audiences pay relatively less attention to the support with more information.

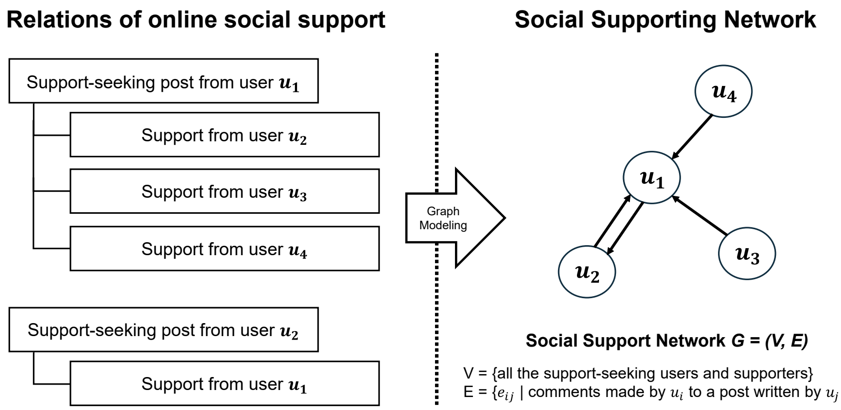


Fig. 4. The process of building social support network from the support-seeking posts and their replied comments.

5 ANALYSIS ON SOCIAL SUPPORTING NETWORK

In this section, we further conduct an in-depth analysis of supporting activities by giving a notion of the social supporting network, whose nodes and edges are the sets of users and supporting comments, respectively. Based on the social supporting network, we explore the behavioral and content characteristics of the heavily-supportive users and the attractive support-seekers who elicit much support from a number of supporters, which answers RQ2. To respond to RQ3, we identify structural communities in social supporting networks and measure how users are grouped and what topics are mentioned in the communities.

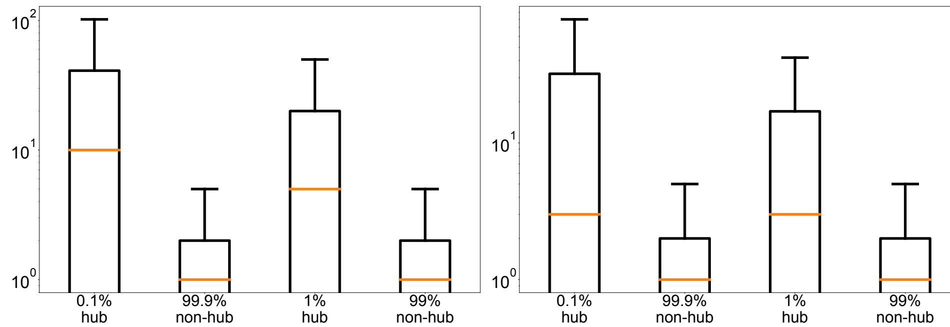
5.1 Building Social Supporting Network

We model the relations of user interaction for online social supports as a social supporting network, illustrated in Figure 4. The social supporting network is a directed graph $G = (V, E)$, where V and E are the sets of the users and the supporting comments, respectively. Each edge connects from node i to node j if user j gives a supporting comment toward the support-seeking posts written by user i. Note that a reciprocal communication of a pair of users (e.g., the edges between u_1 and u_2) can be made when the users replied to each other's posts. Using the collected dataset, we build the social supporting network with 691,426 nodes and 1,962,673 edges. The number and portion of the reciprocal edges are 7,066 and 0.036%, respectively. The number of the strongly connected components, whose every pair of nodes can reach each other, is 119 associated with only 70,073 (10.13%) nodes, implying that the communications among supporters and support-seekers tend to be one-directional and sporadic.

5.2 Analysis on Heavily-Supportive Users and Attractive Support-Seekers

5.2.1 User Selection. To identify heavily-supportive users and attractive support-seekers, we select the top 0.1% and 1% hub users in terms of out-degrees and in-degrees, each of which indicates the supporters who help a number of support-seekers and support-seekers who elicit much support from a number of supporters, respectively. The numbers of top 0.1% (and 1%) users by in-degrees and out-degrees are 693 (6,980) and 697 (6,991), respectively.

We first observe how many users are heavily-supportive and attractive support-seekers simultaneously by measuring Jaccard coefficients between both sets of the in-degree hub and out-degree



(a) In-degrees of top users by out-degrees

(b) Out-degrees of top users by in-degrees

Fig. 5. The distributions of in-degrees of hub and non-hub users by out-degrees and the out-degrees of in-degree hub and non-hub users.

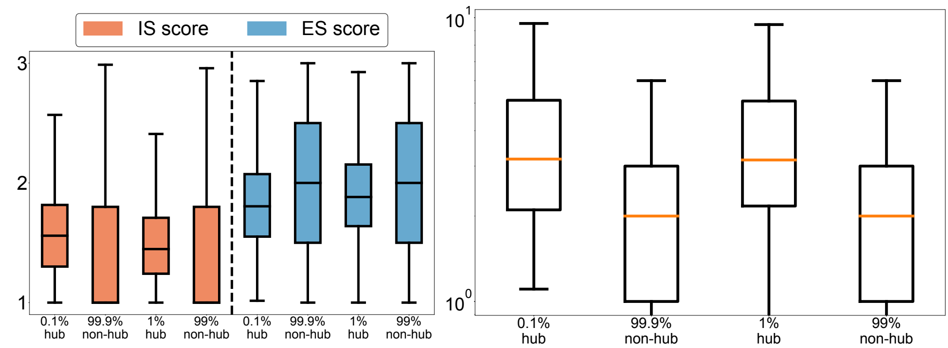
hub users, essentially defined as:

$$J_k = \frac{|U_{in}^k \cap U_{out}^k|}{|U_{in}^k \cup U_{out}^k|} \tag{2}$$

where U_{in}^k and U_{out}^k indicate the top k% users ranked by in-degrees and out-degrees, respectively. Note that the Jaccard coefficient reaches 1 when all the users are in both sets and 0 for the opposite case. The values of $J_{0.1}$ and J_1 are 0.03 and 0.08, respectively, in our case, which indicates that heavily-supportive users and attractive support-seekers are rarely overlapping, instead are distinct.

Figure 5 illustrates the distribution of in-degrees of top 0.1% and 1% users by out-degrees and the out-degrees of top 0.1% and 1% users by in-degrees. In general, the hub users have the significantly higher in-degrees and out-degrees than non-hub users (p<.001); the median of the in-degrees of both top 0.1% and 1% users by out-degrees are 10 and 9, respectively, implying that the heavily-supportive users are also likely to receive more supports from others. Similarly, the out-degrees of the top 0.1% and 1% users by in-degrees have more than two out-degrees while the ones of non-hub users are around 1, indicating that the users who received support from many others tend to also give support to more users. This is in line with the prior study [12], which revealed that the users receiving support tend more to be engaged in social support in online communities. Our result goes one step further – not only the users who receive much support are more likely to participate in social support by giving support, but also the heavily-supportive users tend to have an experience of being supported.

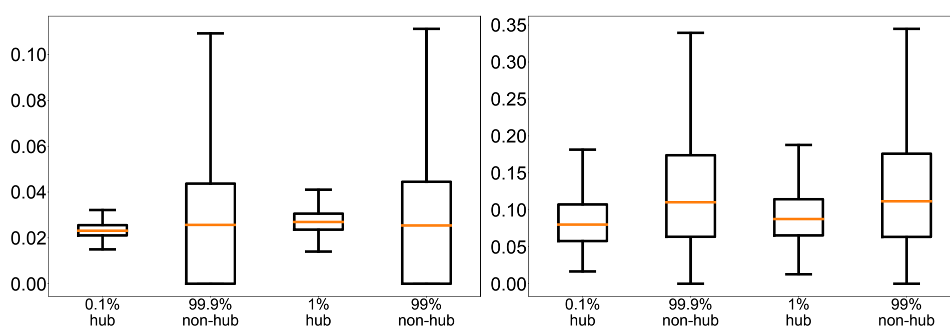
5.2.2 Characteristics of Heavily-Supportive Users. We next investigate what support is provided by heavily-supportive users and how. To this end, we first measure the distribution of the average IS/ES scores and the voted scores of the supporting comments made by the top users by out-degrees, shown in Figure 6. The boxes at the left and right sides (colored by orange and blue, respectively) of Figure 6a indicate the distributions of average IS and ES scores, respectively. Interestingly, the average IS scores of top 0.1% and 1% users are much higher than the ones of non-hub users while top users show relatively lower ES scores than non-hub users, showing the difference of supporting resources between heavily-supportive users and others. That is, the support by these users tends more to be informational although more emotional support is given from the majority of the supporters. Note that all the p-values of t-tests are less than 0.001, which confirms the significance of the results.



(a) Average IS and ES scores

(b) Average voted scores

Fig. 6. The distributions of the average IS/ES scores and the voted scores of the supporting comments made by the out-degree hub users.



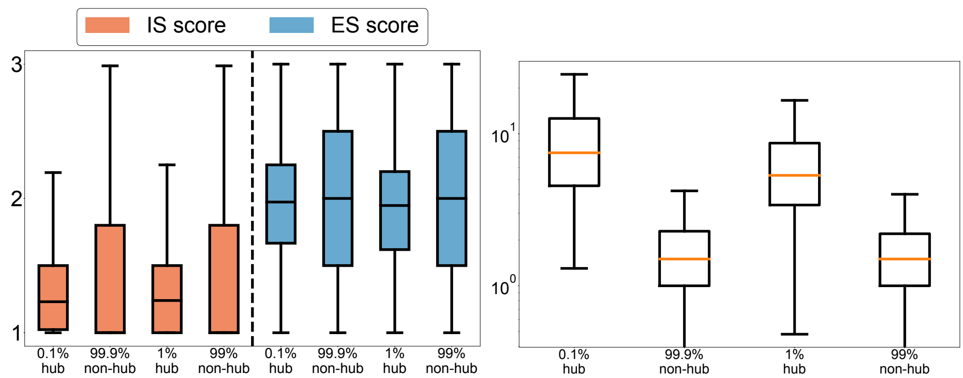
(a) Titles of Participating Posts

(b) Comment Texts

Fig. 7. The distributions of the average TF-IDF similarity scores of the titles of the participating posts and the comment texts made by the hub and non-hub users.

Furthermore, when we look at Figure 6b, top 0.1% and 1% users by out-degrees show higher voted scores on average (p<.005 and p<.065, respectively). Considering the result from Figure 3b, which shows the negative correlation between informational scores and the number of the voted scores, this result implies an important implication on informational support in the social support system; the informational support given by heavily-supportive users tends more to have important features that elicit empathy from community audiences while the ordinary informational support is unlikely to be attractive.

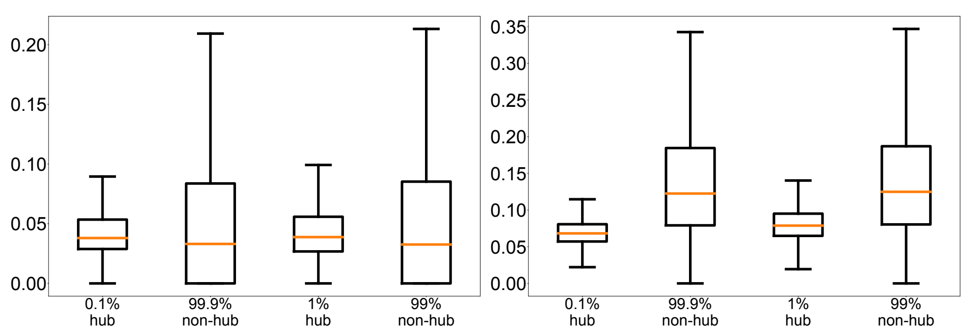
We further find the distribution of the average TF-IDF similarity scores of the titles of the participating posts and the comments. As shown in Figure 7a, the boxes of top 0.1% and 1% users are smaller than non-hub users, indicating that the diversity of the participating topics is different across individual supporters. When we compare the TF-IDF scores for 0.1% and 1% hub, the similarity scores for 0.1% users are lower than the ones for 1% users (p<.001). Furthermore, the tendency is similarly observed in the case of the comment texts (p<.001), as described in Figure 7b. These results imply that the supporters who help extremely many users are more willing to participate in the posts with diverse topics and reply using more diverse words.



(a) Average IS and ES scores

(b) Average voted scores

Fig. 8. The distributions of the average IS/ES scores and the voted scores for the supporting comments.



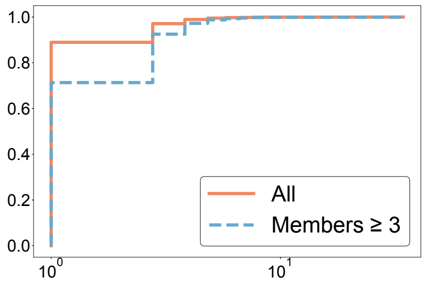
(a) Titles of Uploaded Post

(b) Texts of Received Comments

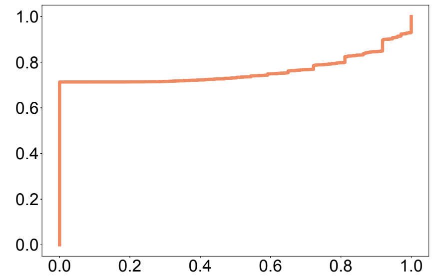
Fig. 9. The distributions of the TF-IDF scores calculated for the titles of the posts uploaded by in-degree hub users and the texts of the comments given to the hub users.

5.2.3 Characteristics of Attractive Support-Seekers. We now focus on in-degree hub users, who elicit supporting comments from a large number of supporters. Similar to the analysis on heavily-supportive users, we measure the distributions of the average IS/ES scores and the voted scores of the supporting comments given to in-degree hub and non-hub users, which are illustrated in Figure 8. The median values of the average IS scores for the comments given to top 0.1% and 1% in-degree hub users are higher than the ones for non-hub users ($p<.001$) while ES scores are similar across all the users ($p>.08$, indicating no significant difference). These results imply that relatively more information is provided to in-degree hub users but the amounts of emotional support given to hub and non-hub users are similar. Note that average voted scores of the supporting comments to both types are significantly different ($p<.001$), as shown in Figure 8b, which means that community audiences also tend more to express empathy to the support given to in-degree hub users.

We finally investigate how similar (i) the posts uploaded by in-degree hub users are and (ii) the supporting comments given to the hub users, by calculating TF-IDF scores for the titles of the posts and the texts of the received comments, respectively. As shown in Figure 9, the median similarity scores of the post titles for top 0.1% and 1% in-degree hub users are higher although the range of the boxes is larger for non-hub users. On the other hand, the tendency is reversed in the case of the



(a) The number of subreddit by community



(b) Entropy of subreddit by community with three or more users

Fig. 10. The distributions of the number of subreddit by all communities and communities with three or more users and entropy of subreddit by communities with three or more people.

comment texts with more significance. Putting these results together, we conjecture that attractive support-seekers constantly upload multiple posts with similar topics, which resulted in receiving more support with diverse topics. This can be a situation where the amount or quality of the given support to a support-seeker is not enough so the support-seeker keeps seeking better support by uploading multiple posts.

5.3 Community Analysis

We next investigate how supporters and support-seekers gather in social support networks by analyzing structural communities, the groups of users who more intensively communicate with other members than out-of-community users. To identify the communities, we use the Louvain method [6], a well-known community detection algorithm that maximizes the ratio of the number of edges within communities (i.e., modularity) to that of edges between communities. Using the method, we identified 21,370 communities in social support network. Note that four giant communities have more than 117,700 members, whereas 13,141 (61.5%) communities have only two members.

5.3.1 Communities with Subreddits. On the identified communities, we investigate what makes the users belong to the same community. To this end, we first compute the subreddit counts and the subreddit entropy for the in-community edges (i.e., the supporting comments among the members in the same community) of each community. Here, we compute the normalized entropy for a community c, calculated as follows:

$$SubredditEntropy(c) = -\sum_{i=1}^{S_c} \frac{p_i^c \ln p_i^c}{\ln S_c} \quad (3)$$

where S_c and p_i^c are the subreddit counts and the portion of the in-community edges made in i-th subreddit, respectively. Note that the value of the subreddit entropy is in a range of [0, 1], reaches 0 if all the comments among members are generated in a subreddit and 1 if the number of comments are equally shared across subreddits.

Figure 10a shows the distributions of the subreddit counts. The subreddit counts of around 89% of all communities are 1. Even though we only consider the communities whose numbers of the members are greater than two, the high portion 71.2% of the communities is involved in a single subreddit, meaning that users in the same subreddit tend more to gather and form a sub-community of subreddit. On the other hand, the communications in 102 communities are across more than

5 subreddits and the highest subreddit count of a community is 47, showing the existence of the cross-subreddit communities for online social support.

Figure 10b tells us that the values of the subreddit entropy show bimodal distribution. Note that we only show the distribution of the communities whose number of the members are more than two. While the entropy values of the dominant communities engaged in only a subreddit are trivially 0, 20.22% communities show the entropy values within a range of [0.8, 1.0], meaning that supporters and support-seekers intensively communicates across multiple subreddits although most communications are made in a subreddit.

5.3.2 Topics in Community. We further analyze what topics are mentioned within user groups (i.e., structural communities) using Latent Dirichlet Allocation (LDA) [5], a popular method of topic modeling, also used in related research in social media [10, 24, 69], that computes the word sets of each topic. To this end, we first select top 0.5% communities in terms of the number of members, where more than 26 post-comment pairs are created in each community, and then perform topic modeling on the titles of the support-seeking posts shared in each community. In particular, we tokenized a post title and removed stopwords and punctuations for preprocessing. Note that we not only use English stopwords provided by Natural Language ToolKIt (NLTK)[3], but also add the non-topical words that appear in most topics without distinction, which includes 'im', 'going', 'would', 'really', 'another', 'everything', 'oh', and 'ever'.

Due to the heterogeneity of the amount of user interactions (i.e., support-seeking posts and supporting comments), we set the different number of topics for LDA to each community. To this end, we first tried to adopt the elbow method, a well-known method that observes the changes of both the coherence and perplexity with increasing the number of topics and determine the number of topics at the *elbow* point of either rapid increase of coherence or sudden drop of perplexity. In particular, we measured the trends of the perplexity and coherence with changing the number of topics from 5 to 30, and then decided the number of topics at the elbow point for individual communities. Unfortunately, this method is unlikely to divide topics properly – the set of the words across different topics are highly overlapped, indicating that the number of topics are too high. Finally, we reduce the number of topics based on the count of community members by an empirical manner. In particular, the numbers of the topics are set to 10, 5, 4, and 3 for the communities whose member counts are more than 10 K, from 1 K to 10 K, from 100 to 1 K, and less than 100, respectively.

We then measure whether and how topics of communities differs by visualizing the keyword vectors of individual communities based on Word2Vec [42] and t-distributed Stochastic Neighbor Embedding (t-SNE) method [29]. To compute a keyword vector of a community, we first extract top 10 keywords by the weights for each topic, and then put each keyword into the pre-trained Word2Vec model that represents a given word as a numeric vector with n-dimension. Here, we employed a Continuous Bag-Of-Word (CBOW) model whose epoch counts, vector size (dimension), min count, and size of context window are set to 5, 100, 1, and 5, respectively. The keyword vector of a community is then obtained by calculating element-wise mean of all the word vectors of a community, whose dimension is then reduced as 2 by t-SNE method for visualization.

Figure 11 shows the results of t-SNE for keyword vectors of top 0.5% communities. The location of each point is calculated from the keyword vector so that the communities whose topic vectors are similar tend to be closer. Note that the coordinate of a point does not provide any information for topics. The sizes and colors of the points vary based on the number of members; large communities are represented as large purple circles while smaller cyan circles indicate the communities with fewer members. Interestingly, the topic vectors of the communities are spread while the large

[3]https://www.nltk.org/

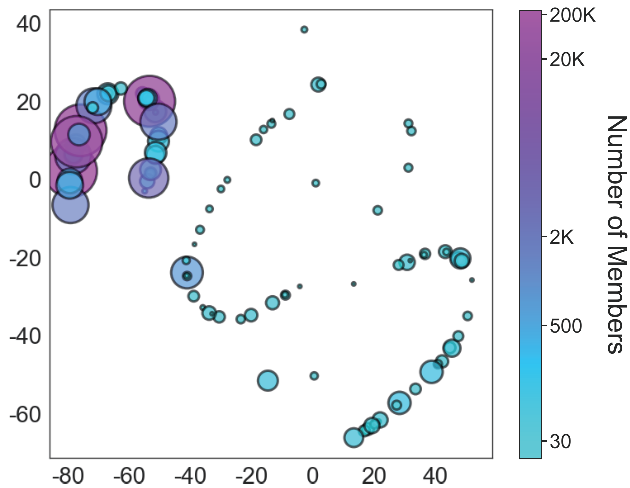


Fig. 11. t-SNE plot of top 0.5% communities in terms of the number of members.

communities tend to be placed on the left-center, implying that supporters and support-seekers are grouped by different topics while the topics of large communities are likely to be similar.

Table 2. Top 10 keywords of each topic for the post titles and the comment texts in the giant community that consists of 135,933 members engaged in 47 subreddits. We bold the distinct keywords of each topic. A detailed version, including the weights of each keyword is provided in Table 6 in Appendix.

	Post Titles
P1	**advice**, today, need, therapist, end, never, daydream, anything, birthday, read
P2	**ocd**, hate, **job**, question, deal, think, new, looking, interview, **diagnosed**
P3	need, feeling, normal, tell, therapy, guy, head, thought, **suicidal**, better
P4	**anxiety**, social, thought, **intrusive**, life, **symptom**, tired, **relationship**, worst, **physical**
P5	**friend**, **scared**, **depressed**, kill, best, **love**, **disorder**, **losing**, done, **lonely**
P6	one, talk, take, could, wish, person, go, care, **depression**, bad
P7	**anxiety**, **attack**, **panic**, this, **heart**, due, weird, cause, anybody, call
P8	**work**, **anxious**, getting, **anxiety**, alone, wrong, **school**, **family**, back, always
P9	**mental**, **health**, stop, fuck, even, **illness**, think, sure, **compulsion**, much
P10	time, **day**, **year**, life, first, every, daydreaming, right, **sleep**, **week**
	Comment Texts
C1	hand, one, even, also, lol, look, **driving**, **car**, face, think
C2	year, time, day, **work**, **job**, **school**, back, week, go, month
C3	**ocd**, thought, post, **tip**, **intrusive**, **anxiety**, way, **compulsion**, how, deal
C4	**anxiety**, **attack**, **panic**, time, **anxious**, **yes**, feeling, **sleep**, day, **heart**
C5	you, hope, **better**, this, **sorry**, life, good, you're, love, **happy**
C6	time, day, one, **friend**, even, go, talk, always, every, **alone**
C7	**try**, **find**, good, also, something, maybe, work, time, need, **book**
C8	**anxiety**, **mental**, **health**, **doctor**, **medication**, **depression**, take, therapy, therapist, **med**
C9	**focus**, **time**, take, **body**, **breathing**, **try**, **breath**, mind, go, thought
C10	think, way, life, you, something, time, say, person, one, even

Inspired by the result, we further investigate what topics are dealt within each community by analyzing the topics and their associated keywords of a few communities as a case study. We first observe the top 10 keywords of each topic in the giant community, where 135,933 members are grouped from 47 subreddits, as listed in Table 2. Here, we not only report the LDA results for the post titles but also for the comment texts. Each topic for the post titles and the comment texts is labeled from P1 to P10 and C1 to C10, respectively.

As shown in Table 2, the words relevant to general mental illness/symptoms/disorders (e.g., ocd, anxiety, depression, and compulsion) commonly appear in the topics for both the post titles and the comment texts. Furthermore, the words that may describe the suffering situation or places (e.g., work, job, school, family, and friend) are also mentioned in both the posts and the comments, implying that social support for diverse topics, which covers a broad range of mental health in general, are made in large communities. Note that the keywords of the topics for the post titles and comments texts are different; the words 'diagnosed', 'suicidal', 'disorder', 'losing', 'lonely', which seem to describe the specific symptoms or disorders in mental health, are used in post titles while the words related to informational support (e.g., 'tip', 'doctor', 'medication', 'breathing' and 'med') and emotional support (e.g., 'yes', 'better', 'sorry', and 'happy') appear in only comment text.

Interestingly, there seem to be specific topics mainly mentioned in the community. When we observe the keywords of the topics for post titles, obsessive-compulsive disorder (OCD) associated with job or work, suicide-relevant therapy, and the mental illness related to a heart attack are dealt in P2, P3, and P7, respectively. The anxiety or depression that come from the relationship with friends or lover is mentioned in P4 and P5 while similar illness related to the surrounding environment (e.g., family, school, or work) is grouped in P8. Note that the word 'advice' is mainly used in P1, implying that the requests for advice are exhibited in the post titles. Similarly, the supports are also separately grouped by different ways, based on how to cope with the suffering situation for support-seekers; driving, sleeping, talking, visiting doctors, and breathing are suggested in C1, C4, C6, C8, and C9. Note that the supports expressing empathy are also in a group (C5).

Table 2 also shows the correspondence of the topics between support-seeking posts and supporting comments; there are a few pairs of topics consisting of a topic for post titles and its corresponding topic for comment texts. The example pairs are P2 with C3, P7 with C4, and P5 with C6, whose primary topics seem to rely on OCD, mental illness after a heart attack, and loneliness related to friendship, respectively. Considering that the community consists of a large number of users across multiple subreddits, these results leverage the potential of design implication to provide collective intelligence of social support across subreddits. That is, when a user uploads a support-seeking post for a topic, the service provider can retrieve the most relevant topic of the given post and its corresponding topic of comment texts and then provide information extracted from the comments related to the corresponding topic. For example, the social supports related to C4, which are collected across multiple subreddits, can be given to the support-seeker who wants to be helped for a mental illness after a heart attack. While the information in a single community (e.g., subreddit) is available for the support-seeker at the moment, this method can provide the information collected from cross-communities.

We further investigate what topics are mentioned in the communities with a small number of members. In particular, we perform the same experiments on three communities whose number of members are 1,860, 702, and 35, across 23, 16, and 5 subreddits, respectively, as a case study. Table 3, 4, and 5 show the results of the topic modeling for three communities. Interestingly, the topics of each community are more specific and almost identical to each other. For example, the keywords related to 'bullying' are in all the topics in Table 3 and Trichotillomania, also known as a hair-pulling disorder, seems to be dealt with in the community as described in Table 4. On the other hand, the topics of the community described in Table 5 rely on excoriation disorder, which is

Table 3. Top 10 keywords of each topic for the post titles and the comment texts in the community that consists of 1,860 members engaged in 23 subreddits. The detailed version of this table including the weights of each keyword are provided in Table 7 in Appendix.

	Post Titles
P1	**bully**, suicide, school, **bullying**, go, friend, much, deal, hate, feeling
P2	**bullying**, **bullied**, **bully**, stop, lost, still, parent, say, this, kid
P3	**bullied**, need, think, life, friend, advice, kill, living, day, **bully**
P4	die, **bullying**, **bully**, school, **bullied**, guy, need, suicide, suicidal, person
P5	depression, life, year, **bullying**, depressed, one, friend, end, **bullied**, never
	Comment Texts
C1	you, need, school, work, **bully**, time, one, think, something, way
C2	life, time, even, better, think, one, go, way, say, need
C3	you, way, **bully**, take, school, friend, life, **bullying**, find, one
C4	school, one, time, life, something, you, think, find, year, back
C5	friend, one, life, even, think, time, school, something, talk, way

Table 4. Top 10 keywords of each topic for the post titles and the comment texts in the community that consists of 702 members engaged in 16 subreddits. The detailed version of this table including the weights of each keyword are provided in Table 8 in Appendix.

	Post Titles
P1	**hair**, **pulling**, year, started, old, triggered, stop, first, **picking**, again
P2	**pulling**, **hair**, **eyelash**, day, pubic, advice, time, end, update, work
P3	**trich**, **pull**, **eyebrow**, **head**, year, past, look, much, finally, **hair**
P4	left, **hair**, bad, one, alone, need, making, back, **eyebrow**, stop
	Comment Texts
C1	you, time, take, this, friend, something, even, try, thank, **pull**
C2	**pulling**, **pull**, try, **hair**, time, life, also, think, even, lot
C3	**hair**, **pull**, time, one, **pulling**, good, look, year, **trich**, also
C4	**hair**, **pulling**, **pull**, **trich**, ive, also, lot, keep, **head**, time

Table 5. Top 10 keywords of each topic for the post titles and the comment texts in the community that consists of 35 members engaged in 5 subreddits. The detailed version of this table including the weights of each keyword are provided in Table 9 in Appendix.

	Post Titles
P1	life, **picking**, long, post, complaining, **face**, **nail**, go, keep, right
P2	**picking**, **skin**, **pick**, **scar**, **picked**, off, **scalp**, ive, rid, long
P3	**scalp**, **picking**, addicted, tip, right, anxiety, **hair**, loss, damage, report
	Comment Texts
C1	still, friend, **scar**, **skin**, one, much, **picking**, **pick**, you, happy
C2	**scalp**, **picking**, **pick**, **dandruff**, **hair**, find, much, **psoriasis**, time, one
C3	you, fluoxetine, time, find, need, take, better, is, within, work

also referred to as chronic skin-picking. Considering that the giant community deals with multiple topics, we conjecture that communities with a small number of members are formed in a center of a more specific topic.

6 DISCUSSION

In this paper, we conducted a quantitative analysis of online social support for mental health from a supporter perspective. We now describe the implications and the limitation of this work.

6.1 Social Media as A Space of Emotional Support

The prior work on online social support has revealed that social media play a significant role in providing the spaces for online social support and there exist multiple types of online social support including informational and emotional support [2, 18, 19, 56]. The investigation in Section 4 on the content characteristics of supporting comments has shown that the dominant portions of online social support in mental health are emotional, which goes one step further from the findings of the prior studies. That is, while the prior studies explored what roles and types of online social support exist, we additionally reveal that online social supports for mental health tend more to be emotional. Note that the result is also in line with the work by Yang et al. [67], who found that emotional support is a majority of social support in online cancer communities. Our work demonstrates that the tendency can be extended to online communities for mental health.

Furthermore, by analyzing the relations between the voted scores and IS/ES scores, we showed that the community audiences are more likely to give the agreement to emotional support, indicating online communities for mental health form an active emotional atmosphere. We argue that this result provides an important insight into how to use online social media for social support to support-seekers or clinicians. That is since the atmosphere of online social support tends to be emotional, not only support-seekers can use online social media as a space to receive emotional support, but also clinicians can encourage the users to exploit social media for providing emotional support.

6.2 Leveraging Informational Support from Heavily-Supportive Users

Although informational support is unlikely to be attractive in online communities in our context, we found that the heavily-supportive users, who support several support-seekers, tend to provide informational support, which attracts more upvotes from community audiences (Section 5.2.2). This result implies that valuable informational support tends to be delivered to heavily-supportive users, which provides an important insight from both theoretical and practical perspectives. From the theoretical perspective, the result opens a few research questions such as “what are distinct features of the informational support by heavily-supportive users?” or “what information is importantly considered in online communities for social support?”, which has been unexplored yet, but can give implications for a comprehensive understanding of online social support. From a practical perspective, this result can be used to leverage informational support on online platforms. For example, the service provider can either recommend heavily-supportive users to support-seekers who want to receive informational support. Moreover, there have been a few studies on improving social support by educating the users with professional guidance, which describes how to deliver the experience or knowledge to support-seekers [12, 43, 50]. Our findings can thus complement these prior studies. That is, a service provider can select heavily-supportive users, who are believed to provide more useful informational support and request to learn professional guidance on specific supporting expressions, to improve the quality of the informational support.

6.3 Consistent Uploads of Similar Posts by Attractive Support-Seekers

In Section 5.2.3, it is demonstrated that in-degree hub users in social support network, who are attractive support-seekers that elicit supporting comments from a number of users, are likely to upload more posts on similar topics. These results can be interpreted as the behavioral characteristics of support-seekers reported in a few prior studies [12, 44, 55]. For example, the support-seekers may be more engaged in online communities, by not only giving support to other support-seekers but also uploading more posts with subsequent topics to be supported. It is also possible that the support-seekers keep seeking social support until being satisfied. A more in-depth analysis of the characteristics of behaviors and content of in-degree hub users can provide a better understanding of the support-seekers' behaviors triggered after support is provided.

6.4 Recommendation based on Structural Communities in Social Supporting Network

The topic analysis on structural communities in Section 5.3 demonstrates that there are a few giant communities consisting of several supporters and support-seekers across multiple subreddits, and a broad range of topics related to mental health is covered in these communities. The result implies that giant communities can be used as effective sources for content or supporter recommendation. On the current social support platforms such as online mental health forums[4], support-seekers have to find a topical community (e.g., subreddit) whose topics match their interests with the expectation that the community members can provide suitable support. More importantly, the support-seekers should spend more time on learning and following not only the explicit rules of the communities to avoid the uploaded post being deleted or transferred by moderators but also implicit *community norms* (e.g., linguistic style) to attract desired supports, as reported in prior work [11, 49]. Provided that the topic of a support-seeking post is accessible, a service provider can retrieve its relevant topics and suggests the associated supports extracted from a giant community, which can increase the opportunity to provide rich support for the given topic from the collective intelligence contributed by the supporters from multiple subreddits.

On the other hand, our investigation of the structural communities with a small number of members demonstrated that small communities dealing with more specific topics such as Trichotillomania or excoriation disorder are successfully identifiable, which can improve existing peer recommendation systems by *community-based* peer suggestion. That is, for a given support-seeking post, the recommendation system first extracts the topical keywords of the given post, finds the communities dealing with similar topics based on the results of the topic modeling, and finally recommends the members of the community as the candidate supporters to the user who uploaded the post. Note that this method can complementarily work with the proposed peer recommendation systems proposed in prior work [27, 68], which suggests peers based on similarity of interests or demographics [27] or social connection [68].

6.5 Limitations

Despite the findings that give insights into understanding online social support for mental health, we note the limitations of this paper. First, our analysis relies on a BERT-based model to score IS/ES of each support in the newly-collected dataset in 2020. In particular, the proposed model learns from the ground-truth dataset consisting of the support-seeking posts and the replied comments with IS/ES scores assigned by two experts manually. Although we validate the proposed model using the ground-truth dataset, which outperforms the state-of-the-art model, evaluating the proposed model with the newly-labeled subset of the collected dataset can be more accurate. However, we expect

[4]It is also called online mental health communities (OMHCs) [49]. To avoid confusion with the communities extracted from the social supporting network in this paper, we noted OMHCs as forums.

the validation difference is not much different with the assumption that the criteria for assigning scores changes over time. Second, we considered only English-written posts and comments created on Reddit. Thus, we caution against generalizing the methodology and findings to other languages or platforms. Third, we observed only the first-level comments like the prior study [56] to consider the social direct support to the original post. A further investigation with supporting comments at the deeper levels may provide different insights.

7 CONCLUSION

In this paper, we investigated online social support from the perspectives of the supporters, who have informational or emotional resources that may affect support-seekers either positively or negatively. In particular, we collected and analyzed a large scale of the supporting comments and their target posts from 55 mental health communities in Reddit, with IS and ES scores assigned by a developed BERT-based deep learning model trained with the ground-truth dataset. From the dataset, we measured the characteristics of the supporters from the behavioral and content perspectives, revealing that the supporters tend to provide emotional support and the users in online communities form an emotional atmosphere for social supports. After modeling the relations between the supports and the support-seekers as "social supporting network", we selected and analyzed top (hub) users by out-degrees and in-degrees in the social supporting network, which demonstrated that heavily-supportive users are more likely to give informational support with diverse content while the users who attract much support exhibit continuous support-seeking behaviors by upload multiple posts with similar content. Our further analysis on structural communities in social supporting network revealed that the communities with small numbers of members tend to deal with a specific topic while the multiple, but common topics are discussed in the giant communities, which are formed by a number of users from multiple subreddits. We believe that the methodologies, dataset, and findings can not only expose more research questions on online social supports in mental health, but also provide an insight on improving social support in online platforms.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5A8055785).

REFERENCES

[1] Kimberley R. Allison, Pandora Patterson, Daniel Guilbert, Melissa Noke, and Olga Husson. 2021. Logging On, Reaching Out, and Getting By: A Review of Self-Reported Psychosocial Impacts of Online Peer Support for People Impacted by Cancer. *Proceedings of the ACM Hum.-Comput. Interact.* 5, CSCW1, Article 95 (Apr 2021), 35 pages. https://doi.org/10.1145/3449169

[2] Nazanin Andalibi, Oliver L. Haimson, Munmun De Choudhury, and Andrea Forte. 2018. Social Support, Reciprocity, and Anonymity in Responses to Sexual Abuse Disclosures on Social Media. *ACM Trans. Comput.-Hum. Interact.* 25, 5, Article 28 (Oct 2018), 35 pages. https://doi.org/10.1145/3234942

[3] Nazanin Andalibi, Oliver L. Haimson, Munmun De Choudhury, and Andrea Forte. 2016. Understanding Social Media Disclosures of Sexual Abuse Through the Lenses of Support Seeking and Anonymity. 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, 3906–3918. https://doi.org/10.1145/2858036.2858096

[4] Christopher Beaudoin and Chen-Chao Tao. 2007. Benefiting from Social Capital in Online Support Groups: An Empirical Study of Cancer Patients. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society* 10 (Sep 2007), 587–90. https://doi.org/10.1089/cpb.2007.9986

[5] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3 (Mar 2003), 993–1022.

[6] 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 (Oct 2008), P10008. https://doi.org/10.1088/1742-5468/2008/10/p10008

[7] Gerald Caplan. 1974. *Support systems and community mental health: Lectures on concept development.* Behavioral Publications, Pasadena, CA, US. xvii, 267–xvii, 267 pages.

[8] Gerald Caplan. 1981. Mastery of stress: Psychosocial aspects. *The American Journal of Psychiatry* 138, 4 (1981), 413–420. https://doi.org/10.1176/ajp.138.4.413

[9] Jean Caron, Éric Latimer, and Michel Tousignant. 2007. Predictors of psychological distress in low-income populations of Montreal. *Canadian Journal of Public Health* 98, 1 (Jul 2007), S35–S44.

[10] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine* 3, 1 (2020), 1–11.

[11] Stevie Chancellor, Andrea Hu, and Munmun De Choudhury. 2018. Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174240

[12] Yixin Chen and Yang Xu. 2021. Social Support is Contagious: Exploring the Effect of Social Support in Online Mental Health Communities. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 286, 6 pages. https://doi.org/10.1145/3411763.3451644

[13] Daejin Choi, Jinyoung Han, Taejoong Chung, Yong-Yeol Ahn, Byung-Gon Chun, and Ted Taekyoung Kwon. 2015. Characterizing Conversation Patterns in Reddit: From the Perspectives of Content Properties and User Participation Behaviors. In *Proceedings of the 2015 ACM on Conference on Online Social Networks* (Palo Alto, California, USA) *(COSN '15)*. Association for Computing Machinery, New York, NY, USA, 233–243. https://doi.org/10.1145/2817946.2817959

[14] Sidney Cobb. 1976. Social Support as a Moderator of Life Stress. *Psychosomatic Medicine* 38, 5 (1976). https://journals.lww.com/psychosomaticmedicine/Fulltext/1976/09000/Social_Support_as_a_Moderator_of_Life_Stress.3.aspx

[15] Sheldon Cohen and Harry M Hoberman. 1983. Positive events and social supports as buffers of life change stress 1. *Journal of applied social psychology* 13, 2 (1983), 99–125.

[16] Odd Steffen Dalgard, Sven Bj, Kristian Tambs, et al. 1995. Social support, negative life events and mental health. *The British Journal of Psychiatry* 166, 1 (1995), 29–34. https://doi.org/10.1192/bjp.166.1.29

[17] Robert C Davis, Ellen Brickman, and Timothy Baker. 1991. Supportive and unsupportive responses of others to rape victims: Effects on concurrent victim adjustment. *American Journal of Community Psychology* 19, 3 (Jun 1991), 443–451.

[18] Munmun De Choudhury and Sushovan De. 2014. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (May 2014), 71–80. https://ojs.aaai.org/index.php/ICWSM/article/view/14526

[19] Munmun De Choudhury and Emre Kiciman. 2017. The Language of Social Support in Social Media and Its Effect on Suicidal Ideation Risk. *Proceedings of the International AAAI Conference on Web and Social Media* 11, 1 (May 2017), 32–41. https://ojs.aaai.org/index.php/ICWSM/article/view/14891

[20] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 4171–4186.

[21] 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. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–27.

[22] Linda K George, Dan G Blazer, Dana C Hughes, and Nancy Fowler. 1989. Social support and the outcome of major depression. *The British Journal of Psychiatry* 154, 4 (Apr 1989), 478–485.

[23] Daena J. Goldsmith, Virginia M. McDermott, and Stewart C. Alexander. 2000. Helpful, Supportive and Sensitive: Measuring the Evaluation of Enacted Social Support in Personal Relationships. *Journal of Social and Personal Relationships* 17, 3 (2000), 369–391. https://doi.org/10.1177/0265407500173004

[24] Crystal Gong, Koustuv Saha, and Stevie Chancellor. 2021. " The Smartest Decision for My Future": Social Media Reveals Challenges and Stress During Post-College Life Transition. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.

[25] Xinning Gui, Yu Chen, Yubo Kou, Katie Pine, and Yunan Chen. 2017. Investigating Support Seeking from Peers for Pregnancy in Online Health Communities. *Proceedings of the ACM Hum.-Comput. Interact.* 1, CSCW, Article 50 (Dec 2017), 19 pages. https://doi.org/10.1145/3134685

[26] Heidi Hartikainen, Afsaneh Razi, and Pamela Wisniewski. 2021. Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks. *Proceedings of the ACM Hum.-Comput. Interact.* 5, CSCW1, Article 42 (Apr 2021), 31 pages. https://doi.org/10.1145/3449116

[27] Andrea L Hartzler, Megan N Taylor, Albert Park, Troy Griffiths, Uba Backonja, David W McDonald, Sam Wahbeh, Cory Brown, and Wanda Pratt. 2016. Leveraging cues from person-generated health data for peer matching in online communities. *Journal of the American Medical Informatics Association* 23, 3 (2016), 496–507.

[28] Jennifer Hefner and Daniel Eisenberg. 2009. Social support and mental health among college students. *American Journal of Orthopsychiatry* 79, 4 (Oct 2009), 491–499.

[29] Geoffrey Hinton and Sam Roweis. 2002. Stochastic Neighbor Embedding. In *Proceedings of the 15th International Conference on Neural Information Processing Systems (NIPS'02)*. MIT Press, Cambridge, MA, USA, 857–864.

[30] R Kahn and Toni Antonucci. 1980. Convoys Over the Life Course: Attachment Roles and Social Support. *Life Span Development* 3 (Jan 1980), 253 – 267.

[31] Berton H Kaplan, John C Cassel, and Susan Gore. 1977. Social support and health. *Medical care* 15, 5 SUPPL (May 1977), 47–58.

[32] Naveena Karusala, David Odhiambo Seeh, Cyrus Mugo, Brandon Guthrie, Megan A Moreno, Grace John-Stewart, Irene Inwani, Richard Anderson, and Keshet Ronen. 2021. "That Courage to Encourage": Participation and Aspirations in Chat-Based Peer Support for Youth Living with HIV. 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 223, 17 pages. https://doi.org/10.1145/3411764.3445313

[33] Kaylee Payne Kruzan, Natalya N. Bazarova, and Janis Whitlock. 2021. Investigating Self-Injury Support Solicitations and Responses on a Mobile Peer Support Application. *Proceedings of the ACM Hum.-Comput. Interact.* 5, CSCW2, Article 354 (Oct 2021), 23 pages. https://doi.org/10.1145/3479498

[34] Per E Kummervold, Deede Gammon, Svein Bergvik, Jan-Are K Johnsen, Toralf Hasvold, and Jan H Rosenvinge. 2002. Social support in a wired world: use of online mental health forums in Norway. *Nord. J. Psychiatry* 56, 1 (2002), 59–65.

[35] Taisa Kushner and Amit Sharma. 2020. Bursts of Activity: Temporal Patterns of Help-Seeking and Support in Online Mental Health Forums. In *Proceedings of The Web Conference 2020*. Association for Computing Machinery, New York, NY, USA, 2906–2912. https://doi.org/10.1145/3366423.3380056

[36] Brian Lakey and Edward Orehek. 2011. Relational regulation theory: a new approach to explain the link between perceived social support and mental health. *Psychological review* 118, 3 (2011), 482.

[37] Richard L Leavy. 1983. Social support and psychological disorder: A review. *Journal of community psychology* 11, 1 (1983), 3–21.

[38] Kai-Kuen Leung, Ching-Yu Chen, Bee-Horng Lue, and Shih-Tien Hsu. 2007. Social support and family functioning on psychological symptoms in elderly Chinese. *Archives of Gerontology and Geriatrics* 44, 2 (2007), 203–213. https://doi.org/10.1016/j.archger.2006.05.001

[39] Zachary Levonian, Marco Dow, Drew Erikson, Sourojit Ghosh, Hannah Miller Hillberg, Saumik Narayanan, Loren Terveen, and Svetlana Yarosh. 2021. Patterns of Patient and Caregiver Mutual Support Connections in an Online Health Community. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 275 (jan 2021), 46 pages. https://doi.org/10.1145/3434184

[40] Ilya Loshchilov and Frank Hutter. 2019. Decoupled Weight Decay Regularization. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net. https://openreview.net/forum?id=Bkg6RiCqY7

[41] Bradley C. Love, Brittani Crook, Charee Mooney Thompson, Sarah Zaitchik, Jessica Knapp, Leah E. LeFebvre, Barbara L. Jones, Erin Donovan-Kicken, Emily Eargle, and Ruth Rechis. 2012. Exploring Psychosocial Support Online: A Content Analysis of Messages in an Adolescent and Young Adult Cancer Community. *Cyberpsychology, behavior and social networking* 15 10 (2012), 555–9.

[42] Tomás Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1301.3781

[43] Kathleen O'Leary, Stephen M. Schueller, Jacob O. Wobbrock, and Wanda Pratt. 2018. "Suddenly, We Got to Become Therapists for Each Other": Designing Peer Support Chats for Mental Health. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173905

[44] Zhenhui Peng, Xiaojuan Ma, Diyi Yang, Ka Wing Tsang, and Qingyu Guo. 2021. Effects of Support-Seekers' Community Knowledge on Their Expressed Satisfaction with the Received Comments in Mental Health Communities. 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 536, 12 pages. https://doi.org/10.1145/3411764.3445446

[45] Julie Prescott, Terry Hanley, and Katalin Ujhelyi. 2017. Peer Communication in Online Mental Health Forums for Young People: Directional and Nondirectional Support. *JMIR Ment Health* 4, 3 (02 Aug 2017), e29. https://doi.org/10.2196/mental.6921

[46] Yada Pruksachatkun, Sachin R. Pendse, and Amit Sharma. 2019. Moments of Change: Analyzing Peer-Based Cognitive Support in Online Mental Health Forums. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300294

[47] Julian Rappaport. 1993. Narrative Studies, Personal Stories, and Identity Transformation in the Mutual Help Context. *The Journal of Applied Behavioral Science* 29, 2 (1993), 239–256. https://doi.org/10.1177/0021886393292007

[48] Afsaneh Razi, Karla Badillo-Urquiola, and Pamela J. Wisniewski. 2020. Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376400

[49] Koustuv Saha, Sindhu Kiranmai Ernala, Sarmistha Dutta, Eva Sharma, and Munmun De Choudhury. 2020. Understanding moderation in online mental health communities. In *International Conference on Human-Computer Interaction*. Springer, 87–107.

[50] Koustuv Saha and Amit Sharma. 2020. Causal Factors of Effective Psychosocial Outcomes in Online Mental Health Communities. *Proceedings of the International AAAI Conference on Web and Social Media* 14, 1 (May 2020), 590–601. https://ojs.aaai.org/index.php/ICWSM/article/view/7326

[51] Koustuv Saha, Benjamin Sugar, John Torous, Bruno Abrahao, Emre Kıcıman, and Munmun De Choudhury. 2019. A social media study on the effects of psychiatric medication use. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 440–451.

[52] Koustuv Saha, John Torous, Eric D Caine, and Munmun De Choudhury. 2020. Psychosocial Effects of the COVID-19 Pandemic: Large-scale Quasi-Experimental Study on Social Media. *J Med Internet Res* 22, 11 (24 Nov 2020), e22600. https://doi.org/10.2196/22600

[53] Catherine Schaefer, James Coyne, and Richard S. Lazarus. 2004. The health-related functions of social support. *Journal of Behavioral Medicine* 4 (2004), 381–406.

[54] Jaclyn Schildkraut, Evelyn S. Sokolowski, and John Nicoletti. 2021. The survivor network: The role of shared experiences in mass shootings recovery. *Victims & Offenders* 16, 1 (2021), 20–49.

[55] Ashish Sharma, Monojit Choudhury, Tim Althoff, and Amit Sharma. 2020. Engagement Patterns of Peer-to-Peer Interactions on Mental Health Platforms. *Proceedings of the International AAAI Conference on Web and Social Media* 14, 1 (May 2020), 614–625. https://ojs.aaai.org/index.php/ICWSM/article/view/7328

[56] Eva Sharma and Munmun De Choudhury. 2018. Mental Health Support and Its Relationship to Linguistic Accommodation in Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174215

[57] Cohen Sheldon and Wills Thomas Ashby. 1985. Stress, social support, and the buffering hypothesis. *Psychological Bulletin* 98, 2 (1985), 310–357.

[58] C. Estelle Smith, Zachary Levonian, Haiwei Ma, Robert Giaquinto, Gemma Lein-Mcdonough, Zixuan Li, Susan O'conner-Von, and Svetlana Yarosh. 2020. "I Cannot Do All of This Alone": Exploring Instrumental and Prayer Support in Online Health Communities. *ACM Trans. Comput.-Hum. Interact.* 27, 5, Article 38 (aug 2020), 41 pages. https://doi.org/10.1145/3402855

[59] Eric Stice, Jennifer Ragan, and Patrick Randall. 2004. Prospective relations between social support and depression: Differential direction of effects for parent and peer support? *Journal of abnormal psychology* 113, 1 (Feb 2004), 155–159.

[60] Peggy A. Thoits. 1982. Conceptual, Methodological, and Theoretical Problems in Studying Social Support as a Buffer Against Life Stress. *Journal of Health and Social Behavior* 23, 2 (1982), 145–159. http://www.jstor.org/stable/2136511

[61] Peggy A. Thoits. 1995. Stress, coping, and social support processes: Where are we? What next? *Journal of health and social behavior* Spec No (1995), 53–79.

[62] Joseph Walther and Saxon Boyd. 2002. Attraction to computer-mediated social support. *Communication Technology and Society: Audience Adoption and Uses* (Jan 2002), 153–188.

[63] Eugene Weinstein. 1965. Stigma: Notes on the Management of Spoiled Identity. Erving Goffman. *Amer. J. Sociology* 70, 5 (1965), 636–636. https://doi.org/10.1086/223949

[64] Steve Whittaker, Ellen Isaacs, and Vicki O'Day. 1997. Widening the Net: Workshop Report on the Theory and Practice of Physical and Network Communities. *SIGCHI Bull.* 29, 3 (Jul 1997), 27–30. https://doi.org/10.1145/264853.264867

[65] Andrew J Winzelberg, Catherine Classen, Georg W Alpers, Heidi Roberts, Cheryl Koopman, Robert E Adams, Heidemarie Ernst, Parvati Dev, and C Barr Taylor. 2003. Evaluation of an internet support group for women with primary breast cancer. *Cancer: Interdisciplinary International Journal of the American Cancer Society* 97, 5 (2003), 1164–1173.

[66] Diyi Yang, Robert Kraut, and John M. Levine. 2017. Commitment of Newcomers and Old-Timers to Online Health Support Communities. 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, 6363–6375. https://doi.org/10.1145/3025453.3026008

[67] Diyi Yang, Robert E. Kraut, Tenbroeck Smith, Elijah Mayfield, and Dan Jurafsky. 2019. Seekers, Providers, Welcomers, and Storytellers: Modeling Social Roles in Online Health Communities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300574

[68] Hangzhou Yang and Huiying Gao. 2021. User recommendation in online health communities using adapted matrix factorization. *Internet Research* 31, 6 (01 Jan 2021), 2190–2218. https://doi.org/10.1108/INTR-09-2020-0501
[69] Jiawei Zhou, Koustuv Saha, Irene Michelle Lopez Carron, Dong Whi Yoo, Catherine R. Deeter, Munmun De Choudhury, and Rosa I. Arriaga. 2022. Veteran Critical Theory as a Lens to Understand Veterans' Needs and Support on Social Media. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 133 (apr 2022), 28 pages. https://doi.org/10.1145/3512980

A RESULTS OF TOPIC ANALYSIS FOR STRUCTURAL COMMUNITIES IN DETAIL

Table 6. Top 10 keywords and their weights of each topic for the post titles and the comment texts in the giant community that consists of 135,933 members engaged in 47 subreddits. We bold the importantly distinct keywords of each topic.

	Post Titles
P1	**advice** (0.0287), today (0.0262), need (0.0241), therapist (0.0162), end (0.0157), never (0.0145), daydream (0.0144), anything (0.0134), birthday (0.0125), read (0.0113)
P2	**ocd** (0.1329), hate (0.0383), **job** (0.0329), question (0.0166), deal (0.0109), think (0.0103), new (0.0084), looking (0.0078), interview (0.0077), **diagnosed** (0.0068)
P3	need (0.0475), feeling (0.0230), normal (0.0210), tell (0.0174), therapy (0.0164), guy (0.0158), head (0.0149), thought (0.0144), **suicidal** (0.0141), better (0.0139)
P4	**anxiety** (0.0859), social (0.0634), thought (0.0519), **intrusive** (0.0228), life (0.0177), **symptom** (0.0153), tired (0.0145), **relationship** (0.0126), worst (0.0116), **physical** (0.0104)
P5	**friend** (0.0523), **scared** (0.0313), **depressed** (0.0202), kill (0.0168), best (0.0145), **love** (0.0133), **disorder** (0.0094), **losing** (0.0093), done (0.0085), **lonely** (0.0082)
P6	one (0.0552), talk (0.0290), take (0.0207), could (0.0195), wish (0.0159), person (0.0136), go (0.0121), care (0.0110), **depression** (0.0090), bad (0.0086)
P7	**anxiety** (0.1755), **attack** (0.0448), **panic** (0.0329), this (0.0155), **heart** (0.0117), due (0.0101), weird (0.0081), cause (0.0080), anybody (0.0079), call (0.0072)
P8	**work** (0.0255), **anxious** (0.0245), getting (0.0188), **anxiety** (0.0188), alone (0.0173), wrong (0.0168), **school** (0.0143), **family** (0.0139), back (0.0126), always (0.0119)
P9	**mental** (0.0520), **health** (0.0401), stop (0.0305), fuck (0.0200), even (0.0148), **illness** (0.0123), think (0.0117), sure (0.0116), **compulsion** (0.0106), much (0.0101)
P10	time (0.0584), **day** (0.0379), **year** (0.0350), life (0.0325), first (0.0284), every (0.0160), daydreaming (0.0156), right (0.0138), **sleep** (0.0137), **week** (0.0129)
	Comment Texts
C1	hand (0.0087), one (0.0075), even (0.0067), also (0.0061), lol (0.0059), look (0.0051), **driving** (0.0049), **car** (0.0048), face (0.0046), think (0.0043)
C2	year (0.0199), time (0.0178), day (0.0149), **work** (0.0141), **job** (0.0137), **school** (0.0104), back (0.0095), week (0.0087), go (0.0083), month (0.0080)
C3	**ocd** (0.0705), thought (0.0612), post (0.0212), **tip** (0.0157), **intrusive** (0.0147), **anxiety** (0.0133), way (0.0125), **compulsion** (0.0120), how (0.0104), deal (0.0091)
C4	**anxiety** (0.0431), **attack** (0.0162), **panic** (0.0158), time (0.0102), **anxious** (0.0096), **yes** (0.0091), feeling (0.0084), **sleep** (0.0079), day (0.0077), **heart** (0.0076)
C5	you (0.0271), hope (0.0170), **better** (0.0150), this (0.0146), **sorry** (0.0141), life (0.0130), good (0.0114), you're (0.0108), love (0.0106), **happy** (0.0102)
C6	time (0.0165), day (0.0150), one (0.0123), **friend** (0.0108), even (0.0101), go (0.0094), talk (0.0072), always (0.0070), every (0.0067), **alone** (0.0058)
C7	**try** (0.0154), **find** (0.0135), good (0.0132), also (0.0117), something (0.0106), maybe (0.0101), work (0.0094), time (0.0088), need (0.0080), **book** (0.0074)
C8	**anxiety** (0.0169), **mental** (0.0167), **health** (0.0140), **doctor** (0.0133), **medication** (0.0127), **depression** (0.0106), take (0.0095), therapy (0.0092), therapist (0.0084), **med** (0.0082)
C9	**focus** (0.0141), **time** (0.0132), take (0.0131), **body** (0.0114), **breathing** (0.0113), **try** (0.0111), **breath** (0.0109), mind (0.0102), go (0.0094), thought (0.0091)
C10	think (0.0154), way (0.0103), life (0.0102), you (0.0092), something (0.0083), time (0.0081), say (0.0078), person (0.0075), one (0.0073), even (0.0071)

Table 7. Top 10 keywords and their weights of each topic for the post titles and the comment texts in the community that consists of 1,860 members engaged in 23 subreddits.

	Post Titles
P1	**bully** (0.0263), suicide (0.0126), school (0.0118), **bullying** (0.0104), go (0.0104), friend (0.0103), much (0.0098), deal (0.0092), hate (0.0091), feeling (0.0079)
P2	**bullying** (0.0359), **bullied** (0.0153), **bully** (0.0108), stop (0.0103), lost (0.0096), still (0.0093), parent (0.0079), say (0.0072), this (0.0071), kid (0.0070)
P3	**bullied** (0.0376), need (0.0312), think (0.0237), life (0.0217), friend (0.0185), advice (0.0144), kill (0.0136), living (0.0086), day (0.0084), **bully** (0.0084)
P4	die (0.0270), **bullying** (0.0243), **bully** (0.0215), school (0.0143), **bullied** (0.0131), guy (0.0131), need (0.0127), suicide (0.0120), suicidal (0.0109), person (0.0089)
P5	depression (0.0299), life (0.0173), year (0.0153), **bullying** (0.0147), depressed (0.0139), one (0.0132), friend (0.0125), end (0.0117), **bullied** (0.0114), never (0.0103)
	Comment Texts
C1	you (0.0150), need (0.0078), school (0.0070), work (0.0066), **bully** (0.0059), time (0.0057), one (0.0055), think (0.0052), something (0.0052), way (0.0051)
C2	life (0.0107), time (0.0100), even (0.0085), better (0.0080), think (0.0075), one (0.0067), go (0.0067), way (0.0063), say (0.0062), need (0.0061)
C3	you (0.0118), way (0.0084), **bully** (0.0079), take (0.0072), school (0.0069), friend (0.0060), life (0.0058), **bullying** (0.0058), find (0.0055), one (0.0049)
C4	school (0.0082), one (0.0065), time (0.0062), life (0.0060), something (0.0059), you (0.0058), think (0.0057), find (0.0051), year (0.0050), back (0.0050)
C5	friend (0.0147), one (0.0076), life (0.0075), even (0.0063), think (0.0062), time (0.0061), school (0.0061), something (0.0061), talk (0.0058), way (0.0054)

Table 8. Top 10 keywords and their weights of each topic for the post titles and the comment texts in the community that consists of 702 members engaged in 16 subreddits.

	Post Titles
P1	**hair** (0.0368), **pulling** (0.0272), year (0.0135), started (0.0132), old (0.0127), triggered (0.0100), stop (0.0098), first (0.0085), **picking** (0.0083), again (0.0078)
P2	**pulling** (0.0246), **hair** (0.0218), **eyelash** (0.0210), day (0.0153), pubic (0.0125), advice (0.0116), time (0.0112), end (0.0106), update (0.0102), work (0.0078)
P3	**trich** (0.0248), **pull** (0.0193), **eyebrow** (0.0165), **head** (0.0164), year (0.0162), past (0.0148), look (0.0135), much (0.0129), finally (0.0119), **hair** (0.0109)
P4	left (0.0182), **hair** (0.0166), bad (0.0158), one (0.0125), alone (0.0119), need (0.0115), making (0.0113), back (0.0110), **eyebrow** (0.0107), stop (0.0095)
	Comment Texts
C1	you (0.0081), time (0.0073), take (0.0061), this (0.0060), friend (0.0052), something (0.0049), even (0.0048), try (0.0046), thank (0.0045), **pull** (0.0042)
C2	**pulling** (0.0081), **pull** (0.0079), try (0.0072), **hair** (0.0068), time (0.0061), life (0.0058), also (0.0058), think (0.0055), even (0.0046), lot (0.0043)
C3	**hair** (0.0200), **pull** (0.0101), time (0.0090), one (0.0086), **pulling** (0.0077), good (0.0077), look (0.0071), year (0.0069), **trich** (0.0060), also (0.0059)
C4	**hair** (0.0179), **pulling** (0.0102), **pull** (0.0096), **trich** (0.0069), ive (0.0066), also (0.0053), lot (0.0052), keep (0.0049), **head** (0.0045), time (0.0045)

Table 9. Top 10 keywords and their weights of each topic for the post titles and the comment texts in the community that consists of 35 members engaged in 5 subreddits.

	Post Titles
P1	life (0.0197), **picking** (0.0194), long (0.0193), post (0.0192), complaining (0.0192), **face** (0.0111), **nail** (0.0111), go (0.0111), keep (0.0110), right (0.0110)
P2	**picking** (0.0354), **skin** (0.0314), **pick** (0.0240), **scar** (0.0163), **picked** (0.0163), off (0.0163), **scalp** (0.0132), ive (0.0126), rid (0.0126), long (0.0125)
P3	**scalp** (0.0580), **picking** (0.0540), addicted (0.0235), tip (0.0147), right (0.0147), anxiety (0.0104), **hair** (0.0104), loss (0.0103), damage (0.0103), report (0.0103)
	Comment Texts
C1	still (0.0089), friend (0.0088), **scar** (0.0074), **skin** (0.0061), one (0.0061), much (0.0061), **picking** (0.0060), **pick** (0.0060), you (0.0060), happy (0.0047)
C2	**scalp** (0.0194), **picking** (0.0134), **pick** (0.0114), **dandruff** (0.0114), **hair** (0.0074), find (0.0074), much (0.0074), **psoriasis** (0.0074), time (0.0064), one (0.0064)
C3	you (0.0077), fluoxetine (0.0076), time (0.0076), find (0.0066), need (0.0066), take (0.0066), better (0.0066), is (0.0066), within (0.0066), work (0.0056)

Received January 2022; revised July 2022; accepted November 2022