

# Engagement Patterns of Peer-to-Peer Interactions on Mental Health Platforms

Ashish Sharma<sup>1\*</sup> Monojit Choudhury<sup>2</sup> Tim Althoff<sup>1</sup> Amit Sharma<sup>2</sup>

<sup>1</sup>Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, USA

<sup>2</sup>Microsoft Research, Bangalore, India

<sup>1</sup>{ashshar, althoff}@cs.washington.edu <sup>2</sup>{monojitc, amshar}@microsoft.com

## Abstract

Mental illness is a global health problem, but access to mental healthcare resources remain poor worldwide. Online peer-to-peer support platforms attempt to alleviate this fundamental gap by enabling those who struggle with mental illness to provide and receive social support from their peers. However, successful social support requires users to engage with each other and failures may have serious consequences for users in need. Our understanding of engagement patterns on mental health platforms is limited but critical to inform the role, limitations, and design of these platforms. Here, we present a large-scale analysis of engagement patterns of 35 million posts on two popular online mental health platforms, TALKLIFE and REDDIT. Leveraging communication models in human-computer interaction and communication theory, we operationalize a set of four engagement indicators based on attention and interaction. We then propose a generative model to jointly model these indicators of engagement, the output of which is synthesized into a novel set of eleven distinct, interpretable patterns. We demonstrate that this framework of engagement patterns enables informative evaluations and analysis of online support platforms. Specifically, we find that mutual back-and-forth interactions are associated with significantly higher user retention rates on TALKLIFE. Such back-and-forth interactions, in turn, are associated with early response times and the sentiment of posts.

## 1 Introduction

Mental illness is an alarming global health issue with adverse social and economic consequences. Mental illness and related behavioral health problems contribute 13% to the global burden of disease, more than cardiovascular diseases and cancer (Collins et al. 2011). Still, access to mental health care is poor worldwide. Most low-income and middle-income countries have less than one psychiatrist per 100,000 individuals (Rathod et al. 2017). Even in high-income countries like the United States, 60% of counties do not have a single psychiatrist (New-American-Economy Research, 2019).

\*This work was done when the author was a Research Fellow at Microsoft Research, India.

Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Research suggests that for people in distress, connecting and interacting with *peers* can be helpful in coping with mental illness, enhancing mental well-being and developing social integration (Davidson et al. 1999). This form of social support (Kaplan, Cassel, and Gore 1977) through peers can be provided online which has stimulated the design and development of online mental health support platforms.

In recent years, several low-cost and easy-to-access peer-to-peer support platforms, such as TALKLIFE & 7Cups<sup>1</sup>, have provided new pathways for seeking social support and dealing with mental health challenges. These platforms allow interactions between support seekers and peers in a thread-like setting; it starts with a user writing a support seeking post which elicits responses from peers and subsequent interactions between the users. Online platforms have multiple advantages over traditional face-to-face supportive methods: they enable asynchronous conversations by design; they are unrestricted by time, space and geographic boundaries; and they facilitate anonymous disclosures which can be helpful in dealing with the major challenge of stigma associated with mental illness (White and Dorman 2001).

However, for these platforms to be successful at facilitating peer-to-peer support, users need to interact and engage. A user who wants to seek support on the platform (henceforth referred to as *seeker*) needs to interact with a peer who is willing to provide support (henceforth referred to as *peer-supporter*). For example, on TALKLIFE, one third of support-seeking posts by users do not receive any responses at all. Receiving no response or having limited engagement with peers can have serious consequences on a mental health platform with vulnerable users, a number of whom are at risk for self-harm or suicide. Also, as indicated in prior literature, engagement between users is key for ensuring favorable outcomes on these platforms (Van Uden-Kraan et al. 2011), including overcoming cognitive distortion, effective distraction, and empathy (Taylor 2011; Mayshak et al. 2017).

Prior work on engagement between users on support platforms have focused on finding its correlations with several user and platform related characteristics, such as methods of support seeking (Andalibi et al. 2018), support provid-

<sup>1</sup><https://talklife.co/>, <https://www.7cups.com/>

ing (Andalibi and Forte 2018), and self-disclosure (Ernala et al. 2018). However, these works have either been conducted as user studies or studies over small human-annotated datasets, or have made strong assumptions in their characterizations of engagement by overlooking factors such as the degree of interaction between users, that are key to engagement in conversations as noted in communication theory (Bretz and Schmidbauer 1983; Williams, Rice, and Rogers 1988; Sheizaf Rafaeli 1988). Furthermore, none of them attempt to develop a collective sense of engagement; instead, they independently examine various engagement dimensions, such as the number of posts and likes.

**Present Work.** In this paper, we conduct a large-scale study of thread-level engagement patterns of 35 million posts across 8 million threads on TALKLIFE and REDDIT. We take a microscopic view of engagement between users on the platform and focus on engagement at the level of conversational thread. Drawing inspirations from Human-Computer-Interaction (O’Brien and Toms 2008) and Communication theory (Bretz and Schmidbauer 1983; Williams, Rice, and Rogers 1988; Rosengren 1999; Ridley and Avery 1979), we operationalize a set of quantitative *indicators* of thread-level engagement around the notions of attention (the amount of attention received by a thread), and interaction (the nature of interaction between the users in the thread) (Section 4). We demonstrate that no single engagement indicator can fully capture observed engagement dynamics. Therefore, we jointly model multiple engagement indicators and discover interpretable thread-level *engagement patterns*. We design a generative model which learns distinct clusters of engagement patterns as a probability distribution over the joint space of multiple engagement indicators (Section 5). We analyze these clusters to derive a set of 11 novel, interpretable engagement patterns (Section 6).

We demonstrate that our novel framework of engagement patterns enables online support platforms to conduct informative self-evaluations and comparative assessments (Section 7). For example, an analysis of TALKLIFE using the framework informs us that a mutual discourse (back-and-forth interactions; Figure 1c) between seekers and peer-supporters is more important for seeker retention than all other engagement indicators. Such an insight is critical for the platforms; platform designers need to uncover design techniques that enable mutual interactions. Moreover, a comparative analysis using our framework highlights the impact of design differences between TALKLIFE & REDDIT to engagement dynamics between users. We end with a discussion of the limitations and risks of our findings for designing mental health support interventions (Section 8).

## 2 Related Work

Our work builds upon the studies of engagement in online communities, research on social support for mental illness and the design of statistical methods for modeling threads.

### 2.1 Engagement patterns in online platforms

The notion of engagement is a complex amalgamation of varied facets; there is no clear way of defining engage-

ment (Smith et al. 2017). The definitions of engagement are adapted based on its context of usage with the focus being on attention, interaction, and affective experience (O’Brien and Toms 2008). Thus, researchers commonly use various context-specific *markers* or *indicators* of engagement. In a recent work on a sexual abuse subreddit, Andalibi et al. (2018) used the length of the thread as the sole indicator of engagement. They found that users who seek direct support receive more replies than the users who do not. Ernala et al. (2018) studied the effect of responder’s engagement on disclosures of highly stigmatized mental illnesses by users on Twitter. They used number of retweets, favorites & mentions on Twitter as their engagement indicators and found positive correlations with the future intimacy of disclosures. Choudhury et al. (2013) studied the engagement of Twitter users before they are diagnosed with depression and used three attention-related engagement indicators – number of posts, number of replies and retweets in response, and two content-related indicators – number of links shared by the user and number of question-centric posts. There has also been work on qualitatively analyzing engagement on MOOC forums. Mak et al. (2010) in their qualitative framework and analysis of MOOC forums differentiate between threads with long-loops (slow single-user posts) and short-loops (quick multi-user posts).

In this work, we focus on thread-level engagement. Our work builds upon prior research by exploring two new interaction-based indicators of engagement (including a novel indicator of degree of interaction based on Communication theory (Bretz and Schmidbauer 1983; Williams, Rice, and Rogers 1988)) which, to the best of our knowledge, have not yet been used in the context of online social support and have limited research in other domains. Moreover, we jointly model and explore these indicators. The use of new indicators and their joint modeling provides us additional insights on the patterns of engagements and their associations with the user behavior on online support platforms.

### 2.2 Mental illness & online social support

There is a rich body of work on detecting and diagnosing mental illness from posts and activities of users on social media platforms (Twitter (De Choudhury et al. 2013; Coppersmith, Harman, and Dredze 2014; Lin et al. 2014; Tsugawa et al. 2015); Facebook (De Choudhury et al. 2014); Instagram (Reece and Danforth 2017)). A similar line of research is focused on estimating the severity of suicidal ideation and risk among individuals disclosing mental illness (De Choudhury et al. 2016; Benton, Mitchell, and Hovy 2017; Gaur et al. 2019). Researchers have also made efforts in differentiating between the various types of support (e.g. *informational*, *emotional*) provided online (Biyani et al. 2014) and analyzing their effects on suicidal risk (De Choudhury and Kiciman 2017) and mental well-being (Saha and Sharma 2020). Studies analyzing self-disclosure (De Choudhury and De 2014; Yang et al. 2019b), anonymity (De Choudhury and De 2014), reciprocity (Yang et al. 2019b), linguistic accommodation (Sharma and De Choudhury 2018) and cognitive restructuring (Pruksachatkun, Pendse, and Sharma 2019) on

Data Statistics	TALKLIFE	REDDIT
# of Threads	6.4M	1.6M
# of Posts	24.9M	9.6M
# of Users	339.4K	969.7K
Observation Period	May 2012 to Jan 2019	Jan 2015 to Jan 2019

Table 1: Statistics of the two data sources.

online support forums have also been conducted. There has also been work on analyzing the quality of online counseling conversations (Pérez-Rosas et al. 2019) and studying their associations with conversation outcomes (Althoff, Clark, and Leskovec 2016).

Our work is directed towards discovering patterns of effective online mental health support conversations. We focus on engagement in support conversations and develop a novel framework that enables informative evaluations and comparative assessments of mental health platforms.

### 2.3 Modeling of a conversational thread

A large set of online platforms such as Reddit, Twitter, etc. facilitate conversations in a thread-like setting. A user on the platform starts a thread which then elicits responses from other users along with back-and-forth interactions between users. This thread-like structure is typically modeled in a generative manner with the focus being on learning the growth dynamics of the thread (e.g. length) or the structural properties of the threads. Kumar et al. (2010) modeled arrival of posts in a thread by incorporating both time and user of the post in a preferential attachment model. Wang et al. (2012) used a continuous-time model over exposure duration and arrival rates of the posts to explain conversational growth. Backstrom et al. (2013) work on the tasks of length prediction and re-entry prediction of users in a thread using features related to post content, time, link and arrival patterns. They also make a distinction between *focused* (long threads with a lot of comments from a small set of users) and *expansory* (long threads with few comments from large set of users) threads. Recently, Aragón et al. (2017) analyzed the differences between linear and hierarchical threads. Lumbreras et al. (2017) propose a mixture model which learns latent roles of users; growth of the thread is dependent on the role of the users in the thread.

In this paper, we build on this literature by jointly modeling multiple engagement indicators and identifying clusters of engagement patterns. To this end, we design a generative model that discovers the desired meaningful clusters in an unsupervised setting (Section 5).

## 3 Dataset Description

We use conversational threads posted on two of the largest online support platforms as our data sources – TALK-

LIFE (talklife.co) and mental health subreddits on REDDIT (reddit.com).

**TALKLIFE.** Founded in 2012, TALKLIFE is a free peer-to-peer network for mental health support. It enables people in distress to have interactions with other peers on the platform. One of the primary ways of having these interactions is using conversational threads which is the focus of our study. A *conversational thread* or simply a *thread* on TALKLIFE is characterized by a user initially authoring a post, typically seeking direct (e.g. *I am struggling with thoughts of self-harm, someone please help*) or indirect (e.g. *Life is like a miserable hassle!*) mental health support; the post then receives (un)supportive responses from the peers on the platform, sometimes leading to back-and-forth conversations between the users. We call the user who authors the first post to start the thread the *seeker* and we call the users who post responses to the thread *peer-supporters*. Note that the notions of seeker and peer-supporter are specific to a thread; a seeker may be a peer-supporter in a different thread. An alternative way of classifying users could be based on their time-aggregated platform activities (Yang et al. 2019a), which is beyond the scope of this work.

**Mental Health Subreddits.** REDDIT is another popular online platform hosting conversational threads. It consists of a large number of sub-communities called *subreddits*, each dedicated to a particular topic. We use threads posts on 55 mental health focused subreddits (list compiled by Sharma et al. (2018)). We accessed the archive of reddit threads hosted on Google BigQuery spanning 2015 to 2019.

**TALKLIFE vs. REDDIT.** REDDIT and TALKLIFE have a key difference in design. REDDIT has topically-focused sub-communities which allows users to subscribe to topics they are most interested in. For example, users dealing with post traumatic stress disorder may only join *r/ptsd*. As we show later, this difference may have major implications on conversational behavior and engagement dynamics on the two platforms (Section 7.1). In addition, only a small part of REDDIT is focused on mental health-related interactions (less than 0.1%) whereas all interactions on TALKLIFE are meant to be ideally focused on mental health. On both the platforms, however, mental health support is provided by volunteer peers (usually untrained) and rarely by professionals.

Table 1 summarizes the statistics of the two datasets. We make use of individual posts in the threads, their timestamps, and the IDs of the users who authored those posts. We minimize the use of additional metadata in our modeling that is highly specific to TALKLIFE or REDDIT (e.g. tags of a thread) in order to promote methods & results that could potentially generalize to other platforms.

**Privacy, Ethics and Disclosure.** The TALKLIFE dataset was sourced (with license and consent) from the TALKLIFE platform. All personally identifiable information was removed before analysis. In addition, all work was approved by Microsoft’s Institutional Review Board. *This work does not make any treatment recommendations or diagnostic claims.*

**Data Access.** The entire REDDIT dataset used in this pa-

per can be accessed from Google BigQuery<sup>2</sup>. A sample of threads from the TALKLIFE dataset can be viewed online<sup>3</sup> but should be used in accordance with their privacy policies and terms of service<sup>4</sup>.

#### 4 Indicators of Thread-Level Engagement

Engagement cannot be fully captured by any single quantitative measure. Instead, various context-specific *markers* or *indicators* are commonly used in engagement studies. In this paper, we are focused on taking a microscopic view of engagement on the mental health forums; we are interested in operationalizing *thread-level engagement*. We look for indicators in a thread which would determine engagement between the seeker and the peer-supporters of the thread. We also keep our framework independent of the content of the individual posts; the content in mental health support platforms is often sensitive and may present ethical concerns. Finally, we want to have a collective and joint understanding of various dimensions of thread-level engagement (Section 5) and thus, wish to use indicators which are complementary. Such a joint understanding is important for finding meaningful *patterns of engagement* (Section 6).

**Engagement vs. outcomes.** While past work has focused on conversation outcomes (Althoff et al. (2016), Pruksachatkun et al. (2019)), here we focus on patterns of engagement. We note that a more engaging thread may not always be the more helpful thread. Depending on context, there might be instances where a thread which is low in engagement is more helpful to a seeker than a highly engaging thread. Future work should investigate the link between engagement & interaction outcomes.

We divide our engagement indicators into 2 categories – *attention-based* indicators and *interaction-based* indicators. The attention-based indicators quantify the amount of attention received by the thread; interaction-based indicators quantify the nature of interaction between seekers & peer-supporters in the thread.

##### 4.1 Attention-Based Indicators

These indicators quantify the amount of attention a thread receives. We use the following indicators based on attention:

**Thread Length.** The number of posts (seeker posts and replies) in a thread. We aim to differentiate between threads with large (*Long Threads*) and small (*Short Threads*) number of posts. We observe that both TALKLIFE and REDDIT contain a large number of threads of length = 1 (32.43% & 27.53% respectively) along with a lot of short threads; long threads are low in proportion. We use generative modeling to determine appropriate thresholds of short and long threads.

**Peer-Supporters.** The number of peer-supporters who post their replies to a thread. We contrast between threads having different engagement dynamics based on the number of peer-supporters, particularly between threads having individual and group communication dynamics (Rosengren

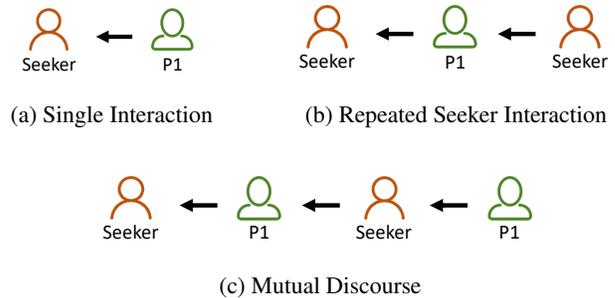


Figure 1: Three types of threads based on degree of interaction between seekers and peer-supporters based on models in communication theory (Bretz and Schmidbauer 1983; Williams, Rice, and Rogers 1988).

1999). If a single peer-supporter responds on a thread, a direct communication takes place between the seeker and the peer-supporter (*Two-Party Threads*). On the other hand, if multiple peer-supporters post on a thread, the communication happens in a group (*Multi-Party Threads*). Moreover, we observe that a lot of threads on both TALKLIFE and REDDIT receive no response at all; they have zero peer-supporters. We follow the terminology used by Ridley & Avery (1979) and call them *Isolated Threads*.

##### 4.2 Interaction-Based Indicators

The second set of indicators capture how seekers & peer-supporters interact with each other. The indicators are:

**Time between Responses.** The time difference between the consecutive posts in a thread. We differentiate between threads with small time between responses (*Quick Threads*) and the threads with large time between responses (*Slow Threads*). Again, instead of manually choosing a threshold we use a generative model to distinguish between the classes.

**Degree of Interaction.** To what extent do the seekers and peer-supporters interact in a thread? We identify three types of threads based on the degree of interaction motivated by communication theory (Figure 1). The first two types of threads are driven by a response from the seeker. The interactive communication theory by Bretz (Bretz and Schmidbauer 1983) differentiates between two interaction mechanisms — interactions in which the sender of a post gets a reply from the receiver but never responds back, and interactions in which the sender gets a reply from the receiver and also responds back. We extend these definitions to the context of threads on online support platforms and define *Single Interaction Threads* and *Repeated Seeker Interaction Threads*. A *Single Interaction Thread* is one in which the seeker of the thread gets a reply from one or multiple peer-supporter(s) but never responds back (Figure 1a). And in a *Repeated Seeker Interaction Thread*, the seeker responds back after a reply from peer-supporter(s) (Figure 1b). We further define a third type of thread, which corresponds to whether a peer-supporter, who had earlier posted on the

<sup>2</sup><https://bit.ly/2WQPosf>

<sup>3</sup><https://web.talklife.co/>

<sup>4</sup><https://www.talklife.co/privacy>, <https://www.talklife.co/terms>

thread, responds back after a response from the seeker (Figure 1c). We call these threads *Mutual Discourse*, a term coined in Williams et al. (Williams, Rice, and Rogers 1988), which relates to higher degrees of interaction between a sender and a receiver.

## 5 Modeling Thread-Level Engagement

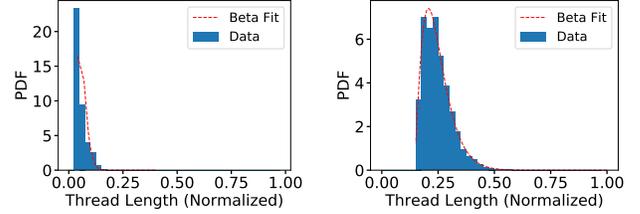
Given the aforementioned indicators of engagement, we aim to systematically discover patterns of engagement in threads. In order to achieve this, we perform the task of modeling thread-level engagement. We start by reasoning for the need of a computational model for discovering the engagement patterns. We then formally describe our modeling assumptions, followed by the generative process of the model and the parameter inference.

**Why do we need a joint understanding of engagement indicators?** Analyzing only a single dimension of engagement for a thread is likely to be insufficient; it won't present a comprehensive view of the engagement dynamics of a thread. Consider the following two dimensions: number of peer-supporters and the degree of interaction (Figure 1). If we are looking at both of them individually, then we might miss out on cases where a seeker is having deep mutual discourse with a lot of peer-supporters; this simultaneous occurrence or interaction of both dimensions may have a stronger effect on engagement and the subsequent conversation outcome, than the occurrence of deep mutual discourse or a lot of peer-supporters, independently. Similarly, two threads with similar number of posts and peer-supporters can have very different engagement dynamics if one has long delays between messages from the peer supporters compared to the other.

**Why do we need a model?** Once the indicators are defined, the potential space of engagement patterns becomes obvious. Specifically, the engagement pattern  $e$  of a thread would lie in the following space generated by the indicators:

$$\begin{aligned}
 e \in & \underbrace{\{\text{Short, Long}\}}_{\text{What is the length?}} \times \underbrace{\{\text{Slow, Quick}\}}_{\text{What is the Time between Responses?}} \\
 & \times \underbrace{\{\text{Isolated, Two-Party, Multi-Party}\}}_{\text{How many Peer-Supporters?}} \\
 & \times \underbrace{\{\text{Single Interaction, Repeated Seeker Interaction, Mutual Discourse}\}}_{\text{What is the degree of interaction between seeker and peer-supporter?}} \quad (1)
 \end{aligned}$$

One might consider manually exploring engagement patterns across all combinations of the four engagement indicators. However, this approach would require extensive manual effort and time, and the need for making arbitrary decisions on thresholds (e.g. long vs. short threads), which would render the approach non-scalable, domain-dependent and subjective. Further, some indicators might be naturally correlated (e.g. thread length and number of peer supporters) in the data, obviating the need for defining certain classes of engagement patterns that are most likely empty or unnatural (e.g. long isolated threads).



(a) Length distribution of threads with  $\delta_{i,j} \geq 100$ ; beta-fit (b) Length distribution of threads which are Multi-Party Mutual Discourse; beta fit

Figure 2: Empirical validation of modeling assumptions.

**Our Model.** We propose a generative model over the set of engagement indicators described in Section 4. The model learns a set of clusters of threads based on a maximum likelihood objective (Equation 2). This type of modeling, as we shall see later, allows us to discover the distinct and interpretable engagement patterns of threads. We discuss the details of our model next.

### 5.1 Modeling Assumptions

We assume that every cluster has its own joint distribution over the engagement indicators of threads; each joint distribution describes a different thread-level engagement pattern. We further assume that each thread is generated from a single engagement cluster. This is helpful in efficient learning of our model parameters primarily due to limited signals available in each thread (Yin and Wang 2014).

Let  $\mathcal{T}$  be the set of threads. A thread  $\mathbf{T}_i \in \mathcal{T}$  consists of an initial post  $p_{i,0}$  by the seeker and a set of  $k - 1$  replies  $p_{i,1}, p_{i,2}, \dots, p_{i,k-1}$  having a total thread length of  $k^5$ . Here, we represent each reply in the thread with a tuple  $p_{i,j} = (u_{i,j}, r_{i,j}, \delta_{i,j})$  where  $u_{i,j}$  is the user of the post,  $\delta_{i,j}$  is the time elapsed since the last post and  $r_{i,j}$  is the role of the user based on the interaction dynamics local to the thread (indicative of first peer-supporter, re-entry of an existing peer-supporter, a new peer-supporter, and seeker's response). We define a set  $\mathcal{R}$  of 4 user roles: **(a) First Peer-Supporter:** user of the first reply of the thread, i.e.,  $j = 1$ ; **(b) New Peer-Supporter:**  $u_{i,j}$  is new to the thread but not the first peer-supporter, i.e.,  $\forall k < j \ u_{i,k} \neq u_{i,j}, j \neq 1$ ; **(c) Existing Peer-Supporter:**  $u_{i,j}$  is a peer-supporter who has interacted with the thread before, i.e.,  $\exists k < j: u_{i,k} = u_{i,j}$  and  $u_{i,k} \neq u_{i,0}$ ; **(d) Seeker:**  $u_{i,j}$  is the seeker, i.e.,  $u_{i,j} = u_{i,0}$ . This categorization of user types based on interaction helps us in accounting for both the number of peer-supporters and the degree of interaction between the seeker and the peer-supporter. We intentionally distinguish between first peer-supporter and new peer-supporter as this allows us to easily differentiate between Two-Party threads and Multi-Party threads; a thread would be Two-Party if there is no new peer-supporter. Moreover, the use of seeker

<sup>5</sup>We combine two consecutive posts by the same user as a pre-processing step. Note that this may result in loss of temporal effects, which is not in the scope of this study.

and existing peer-supporter is helpful in differentiating between Single Interaction, Repeated Seeker Interaction, and Mutual Discourse threads.

**Parametric Assumptions.** We assume that the distributions of thread lengths and time between responses can be well-approximated through Beta distributions. This assumption is based in both theoretical and empirical findings. Theoretically, beta distributions can approximate power laws as well as family of exponential distributions that emerge in natural systems that follow some kind of preferential attachment law (Peruani et al. 2007).

Empirically, thread lengths in online forums usually follow a power-law distribution (Yu et al. 2010). Further, as shown in Figure 2, the length<sup>6</sup> distribution of threads on TALKLIFE—(i) which are potentially slow ( $\delta_{i,j} \geq 100$ ) (Figure 2a), and (ii) in which the seeker has Mutual Discourse with multiple peer-supporters (Multi-Party) (Figure 2b) – are both well-approximated by Beta distributions.

We make similar observations for the potential  $\delta_{i,j}$  clusters. We make use of these observations in our model; we assume that the lengths and the time between responses within each cluster are generated from a Beta distribution.

Moreover, user roles in a thread are assumed to be categorical distributions over the set  $\mathcal{R}$  consisting of the 4 types of roles — First Peer-Supporter, New Peer-Supporter, Existing Peer-Supporter and Seeker, where the distributions themselves are Dirichlet distributed (similar to model assumptions of Latent Dirichlet Allocation (Blei, Ng, and Jordan 2003)).

## 5.2 Generative Process

Let  $\mathcal{E}$  be the set of engagement clusters. A thread  $\mathbf{T}_i \in \mathcal{T}$  is generated as described in Algorithm 1. The engagement distribution  $\theta_{\mathcal{E}}$  is drawn from a Dirichlet distribution with prior  $\alpha_{\mathcal{E}}$  (Line 1). Likewise, for every engagement cluster  $e$ , user-role distributions  $\phi_e^{\mathcal{R}}$  are drawn from a Dirichlet distribution with prior  $\alpha_e^{\mathcal{R}}$  (Line 2-4). For every thread  $\mathbf{T}_i$ , first an engagement cluster  $e$  is chosen from the engagement distribution  $\theta_{\mathcal{E}}$  (Line 6). The thread length  $k$  of  $\mathbf{T}_i$  is sampled from the beta distribution of this engagement cluster  $e$  parameterized by alpha  $\alpha_e^{\mathcal{K}}$  and beta  $\beta_e^{\mathcal{K}}$  (Line 7). Next, the engagement cluster generates the set of replies (Line 8-11). The  $j$ -th reply consists of local user roles  $r_{i,j}$  and time deltas  $\delta_{i,j}$ . Each  $r_{i,j}$  is sampled from the categorical user-role distribution  $\phi_e^{\mathcal{R}}$  (Line 9) and each  $\delta_{i,j}$  is sampled from the beta distribution parameterized by alpha  $\alpha_e^{\delta}$  and beta  $\beta_e^{\delta}$  (Line 10). The likelihood of generating thread  $T_i$  from an engagement cluster  $e$  is given by:

$$p(\mathbf{T}_i|e) \propto \frac{n_e + \alpha_{\mathcal{E}}}{|\mathcal{T}| + |\mathcal{E}| * \alpha_{\mathcal{E}}} * \frac{k^{\alpha_e^{\mathcal{K}} - 1} (1 - k)^{\beta_e^{\mathcal{K}} - 1}}{B(\alpha_e^{\mathcal{K}}, \beta_e^{\mathcal{K}})} * \prod_{P_{i,0}, P_{i,1}, \dots, P_{i,k-1}} \left( \phi_e^{\mathcal{R}}(r_{i,j}) * \frac{\delta_{i,j}^{\alpha_e^{\delta} - 1} (1 - \delta_{i,j})^{\beta_e^{\delta} - 1}}{B(\alpha_e^{\delta}, \beta_e^{\delta})} \right) \quad (2)$$

<sup>6</sup>We use min-max scaling for transforming the lengths and time between responses to  $[0, 1]$  interval, which then can be modeled by beta-distribution.

---

## Algorithm 1 Generative process of our engagement model

---

- 1: Draw engagement distribution  $\theta_{\mathcal{E}} \sim Dir(\alpha_{\mathcal{E}})$
  - 2: **for** each engagement cluster  $e \in \mathcal{E}$  **do**
  - 3:   Draw user-role distribution  $\phi_e^{\mathcal{R}} \sim Dir(\alpha_{\mathcal{R}})$
  - 4: **end for**
  - 5: **for** each thread  $\mathbf{T}_i \in \mathcal{T}$  **do**
  - 6:   Draw an engagement cluster  $e \sim \theta_{\mathcal{E}}$
  - 7:   Draw the thread length  $k \sim Beta(\alpha_e^{\mathcal{K}}, \beta_e^{\mathcal{K}})$
  - 8:   **for** each reply post  $p_{i,j} \in \mathbf{T}_i$  **do**
  - 9:     Draw the user role  $r_{i,j} \sim Multi(\phi_e^{\mathcal{R}})$
  - 10:     Draw the time to reply  $\delta_{i,j} \sim Beta(\alpha_e^{\delta}, \beta_e^{\delta})$
  - 11:   **end for**
  - 12: **end for**
- 

where  $n_e$  is the number of threads in the engagement cluster  $e$ .

Our learning objective for deriving the desired clusters is to maximize this likelihood given a dataset of threads. Note that we do not make use of the Isolated threads while learning the model; these threads are separately identified and integrated with the inferred patterns (Section 6).

## 5.3 Parameter Inference

We use a Gibbs-sampling approach for inferring the Dirichlet distribution. The likelihood of generating a role  $r_{i,j}$  from engagement cluster  $e$  is given by:

$$\phi_e^{\mathcal{R}}(r_{i,j}) = \frac{n_e^{(r_{i,j})} + \alpha_{\mathcal{R}}}{n_e^{(\cdot)} + |\mathcal{R}| * \alpha_{\mathcal{R}}} \quad (3)$$

where  $n_e^{(r_{i,j})}$  is the number of times role  $r_{i,j}$  has been assigned to cluster  $e$ ,  $n_e^{(\cdot)}$  is the marginal count over all roles in  $\mathcal{R}$ . We use method of moments for inferring the Beta distribution parameters —  $\alpha_e^{\mathcal{K}}, \beta_e^{\mathcal{K}}, \alpha_e^{\delta}, \beta_e^{\delta}$  (Wang and McCallum 2006). We initialize the two Dirichlet priors using a commonly used strategy in LDA-based models ( $\alpha_X = 50/|X|, X = \mathcal{E}, \mathcal{R}$ ) (Diao et al. 2012). We optimize on the number of clusters using the popular Elbow method. For both TALKLIFE and REDDIT, we choose the number of clusters as 20 which also gives us the most diverse and interpretable clusters.

## 6 Inferred Engagement Patterns

Using the engagement clusters learned by our generative model, we infer the predominant set of engagement patterns of threads on TALKLIFE and REDDIT. We analyze the distributions which the learned clusters have over the three dimensions of a thread (length, time delta, user roles) —  $Beta(\alpha_e^{\mathcal{K}}, \beta_e^{\mathcal{K}}), Beta(\alpha_e^{\delta}, \beta_e^{\delta})$ , and  $\phi_e^{\mathcal{R}}$ . We do this by simultaneously considering the space over engagement indicators defined in Equation 1. The clusters are manually analyzed and coded by multiple authors in a top-down approach, after which we derive the following *hierarchically organized engagement patterns* (fraction of threads following each pat-

tern is shown in brackets – TALKLIFE & REDDIT respectively):

- **Isolated (32.43% & 27.53%)**
- **Single Interaction (30.57% & 7.64%):**
  - Two-Party (20.30% & 0.08%):
    - (i) Short Slow Two-Party SI (20.30% & 0.08%)
  - Multi-Party (10.27% & 7.56%):
    - (ii) Short Slow Multi-Party SI (10.27% & 7.56%)
- **Repeated Seeker Interaction (18.6% & 21.4%):**
  - Two-Party (4.25% & 5.58%):
    - (iii) Short Slow Two-Party RSI (3.39% & 3.99%)
    - (iv) Short Quick Two-Party RSI (0.86% & 1.59%)
  - Multi-Party (14.35% & 15.82%):
    - (v) Short Slow Multi-Party RSI (1.10% & 12.96%)
    - (vi) Short Quick Multi-Party RSI (13.25% & 2.86%)
- **Mutual Discourse (18.4% & 43.43%):**
  - Two-Party (8.86% & 22.08%):
    - (vii) Short Quick Two-Party MD (8.11% & 21.99%)
    - (viii) Long Quick Two-Party MD (0.75% & 0.09%)
  - Multi-Party (9.54% & 21.35%):
    - (ix) Short Quick Multi-Party MD (6.17% & 17.33%)
    - (x) Long Quick Multi-Party MD (3.37% & 4.02%)

where SI: Single Interaction; RSI: Repeated Seeker Interaction and MD: Mutual Discourse. The engagement patterns are named based on the most likely or dominant set of engagement indicators. We qualitatively evaluate the inferred patterns as described next.

### 6.1 Qualitative evaluation of inferred patterns

We present a qualitative evaluation of the distributions of the inferred engagement patterns over user roles, thread length and time between responses.

**User roles.** Figure 3 shows the distribution of engagement patterns over user roles (marginalizing the patterns along length and time between responses). The Two-Party patterns contain high percentage of peer-supporters who are first to the threads (first peer-supporters) and low percentage of peer-supporters who are new (new peer-supporter);

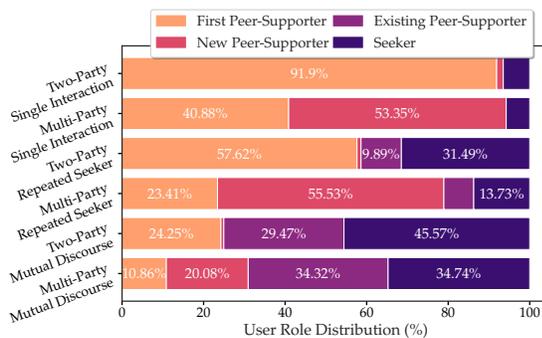


Figure 3: Engagement patterns and user roles.

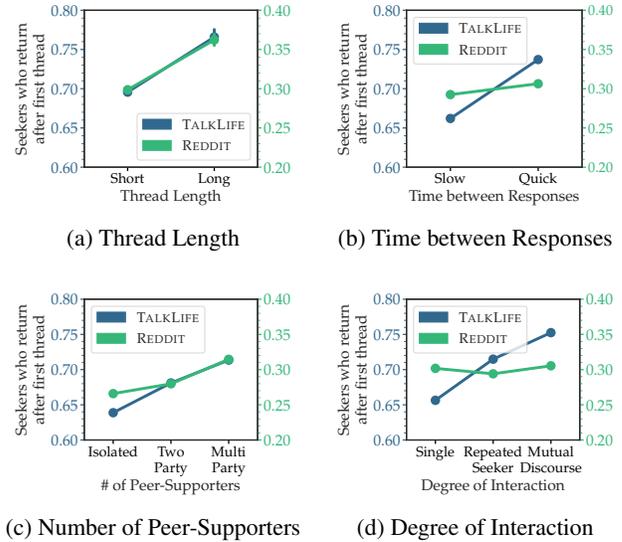


Figure 4: Fraction of seekers who return after their first thread across four engagement indicators. Seekers are more likely to return after higher degrees of engagement. Note that, in this paper, we are only interested in variation of retention likelihood with degrees of engagement and not the absolute values. Error bars throughout the paper are bootstrapped 95% confidence intervals.

this is indicative of presence of only one peer-supporter (the first one) in the thread, hence Two-Party (the other party being the seeker). On the other hand, Multi-Party patterns have higher percentage of peer-supporters who are new, indicative of multiple peer-supporters in the thread. The Single Interaction patterns have low percentage of seekers and existing peer-supporters; these patterns rarely get responses from seekers and existing peer-supporters. Finally, the Mutual Discourse threads have high seeker and existing peer-supporter percentages, indicative of multiple interaction between the seeker and peer-supporter(s) of the thread.

**Thread lengths & time between responses.** Further, we analyze the distributions of the inferred Short, Long, Slow & Quick threads for both TALKLIFE & REDDIT. On TALKLIFE, Short threads have a mean length of 3.9 and a median length of 3. Whereas Long threads have a mean length of 13.5 and a median length of 10. Slow threads have a median time to reply of 7 minutes. Quick threads have a median time to reply of 1 minutes.

On REDDIT, Short threads have a mean length of 3.33 and a median length of 3. Whereas Long threads have a mean length of 23.95 and a median of 19. Slow threads have a median time to reply of 75 minutes. Quick threads have a median time of 16 minutes.

## 7 Implications of Engagement Patterns

We next investigate the implications of our framework of engagement patterns and demonstrate ways in which it can

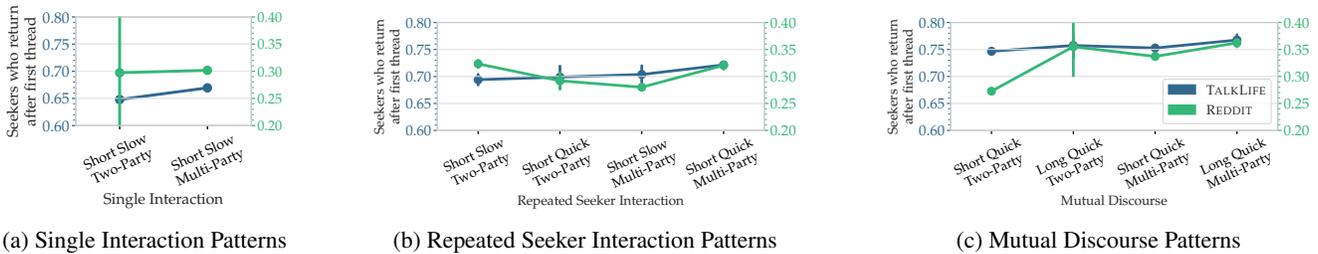


Figure 5: Seeker retention and joint engagement patterns. Degree of Interaction is key to seeker retention; Mutual Discourse is more important for seeker retention than Repeated Seeker Interaction or Single Interaction. Within each interaction degree there is very limited variation between engagement patterns.

guide online support platform evaluations and assessments.

We first exploit our framework for comparing the functioning of TALKLIFE and REDDIT (Section 7.1). Next, we look at the correlations these engagement patterns have with the retention of seekers (Section 7.2) and peer-supporters (Section 7.3) on TALKLIFE and REDDIT. Before we conclude, we investigate the factors which potentially lead to Mutual Discourse between seekers and peer-supporters in a thread (Section 7.4).

## 7.1 Comparative assessment of TALKLIFE & REDDIT

The patterns inferred in Section 6 exhibit an interesting contrast between TALKLIFE and REDDIT. Even though the amount of Isolated threads on the two platforms is comparable, we observe REDDIT to have very few Single Interaction threads involving one peer-supporter (Two-Party; 0.08%). This suggests that, on REDDIT, after an interaction happens on a thread posted by a seeker, it either attracts other users or it spawns a future interaction with the seeker. Also, a much larger fraction of threads on REDDIT result in a Mutual Discourse as compared to TALKLIFE (43.43% vs. 18.4%). Both these differences might be a result of the sub-community nature of REDDIT where users can subscribe to the topics they care about and the topics in which they can provide effective peer-support. For example, a peer-supporter who has dealt with post traumatic stress disorder (PTSD) in the past might be looking for supporting users with PTSD; this can be achieved on the *r/ptsd* subreddit. On TALKLIFE, however, that peer-supporter will need to manually *follow* users having PTSD which may not be intuitive. This indicates that having an organized structure of the threads and users is helpful in making online platforms more engaging. Future work should investigate the relationship of engagement with community structure dynamics.

## 7.2 Seeker Retention on support platforms

Seeker retention, particularly during the initial days of the seeker on the platform, is important for receiving proper support from the peers. We define retention as returning back to the platform after the first thread of interaction by writing another post or response in a different thread (on any of

the mental health subreddits in case of REDDIT). We look into the fraction of seekers who return back when their first threads have a certain engagement pattern.

**Seeker retention increases with higher degrees of engagement.** We examine how seeker retention varies with different degrees of engagement. Figure 4 shows the variation of fraction of seekers who return to the platform after their first threads across the four individual engagement indicators. We observe that seekers are more likely to return after higher degrees of engagement in their first thread; seeker retention is least likely after an Isolated first thread (0.64 & 0.26 respectively for TALKLIFE & REDDIT); Long threads have more seeker retention likelihood than Short threads (0.77 vs. 0.70; 0.36 vs. 0.30); Multi-Party more than Two-Party (0.71 vs. 0.68; 0.31 vs. 0.28). Furthermore, on TALKLIFE, Quick threads have higher seeker retention likelihood than Slow (0.74 vs. 0.66); Mutual Discourse has higher likelihood than Repeated Seeker Interaction and Single Interaction threads (0.75 vs. 0.71 vs. 0.66). This, however, is in contrast with REDDIT, where there is very little variation in seeker retention rates along the indicators of time between responses (0.29 vs. 0.30) and degree of interaction (0.31 vs. 0.29 vs. 0.31).

**Mutual Discourse is more important for seeker retention independent of other engagement indicators.** We next check if the joint presence of the engagement indicators have a role to play on seeker retention. For this, we analyze the variation of seeker retention likelihood with joint engagement patterns. We find that the variation in the likelihood is largely between the three degrees of interaction; the variation is low among the other indicators within a specific degree of interaction (Figure 5). Single Interaction threads have the least likelihood followed by Repeated Seeker Interaction; Mutual Discourse have the highest retention likelihood. The variation is limited within each degree for TALKLIFE; a Mutual Discourse is expected to have the highest seeker retention likelihood independent of whether it is Long or Short, Two-Party or Multi-Party.

To understand how these variations inform us better engagement mechanisms on TALKLIFE, consider a situation where a seeker starts a thread. The likelihood of her return-

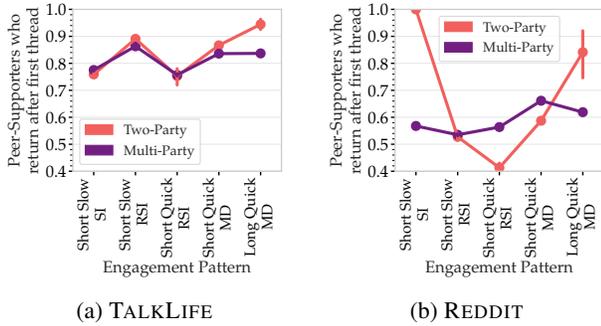


Figure 6: Peer-supporter retention and engagement patterns. On both platforms, the joint presence of Long, Quick and Mutual Discourse indicators elicits differences between Two-Party and Multi-Party patterns.

ing to TALKLIFE if she gets no response is 64%. Receiving a single response from a peer-supporter increases the retention likelihood to 65%. However, if the seeker replies back after the response, likelihood jumps to 69%. At this point, even if a few more peer-supporters post on the thread, the likelihood hovers around 70%. This likelihood gets a major boost and jumps to 75% if one of the existing peer-supporters replies again leading to a Mutual Discourse. This is how important Mutual Discourse is to TALKLIFE, with a gain of 10% over Single Interaction threads and a gain of 5% over Repeated Seeker Interaction threads in terms of seeker retention likelihood.

This provides quantitative evidence that simply connecting users on the platform or trying to have each post get a response may not result in optimal outcomes. Instead, truly mutual interactions were associated with high seeker retention rates. Future work should investigate how to effectively engage seekers and peer-supporters in mutual discourse.

REDDIT, surprisingly, has a low seeker retention likelihood for mutual discourses which are Short, Quick, and Two-Party. This hints towards platform-specific nuances on REDDIT. These are instances of seekers having solitary mutual, back-and-forth conversations with a peer-supporter. A lot of them might have been *throwaway* accounts on REDDIT or the seeker involved might have moved to a different more-suitable subreddit.

### 7.3 Peer-Supporter Retention

Unsurprisingly, peer-supporters are key to a peer-to-peer support platform. Increasing their retention and ensuring that they, as a whole, are providing adequate support on the platform is critical for successful social support. We next investigate retention of peer-supporters and explore the engagement patterns which have an increased likelihood of peer-supporter retention. Similar to our seeker retention analysis, we define retention of a peer-supporter as returning back to the platform after the first thread of interaction (in a new thread). We inspect the correlation of engagement patterns of the threads with peer-support retention likelihood.

**Peer-Supporters return more often if they were the sole supporters.** Analyzing the number of peer-supporters in a

thread, we find that Two-Party threads have more peer-supporter retention likelihood than Multi-Party threads on TALKLIFE (0.81 vs. 0.79;  $p < 0.01^7$ ). The differences between these two threads are much more prominent in the case of Long Quick Mutual Discourse which is also visible for REDDIT (Figure 6; TALKLIFE - 0.95 vs. 0.84; REDDIT - 0.84 vs. 0.62;  $p < 0.001$ ); the greater differences highlight the importance of looking jointly at the indicators. These observations indicate that a peer-supporter is more likely to return if they previously were the only supporter in a thread. This matches the findings in previous work in the context of online crowdfunding, where donors were more likely to return if they were the only donor or one of the very few, presumably due to a stronger sense of personal impact (Althoff et al. (2015)).

**Peer-Supporters who are slower-to-act are more likely to return.** In our analysis of peer-supporter retention in Repeated Seeker Interaction threads, interestingly, we find that Slow threads have higher likelihood of peer-supporter retention than the Quick threads (Figure 6). These are patterns in which only the seeker interacts repeatedly with the thread. This begs the question of why a Slow thread with seeker response and no second response from the peer-supporter (no Mutual Discourse) will have high correlation with peer-supporter retention. For this, we take a deeper look at the threads of the two types (Slow and Quick) and look at the behavior of seeker and the peer-supporter in the thread. We compare the ratios of peer-supporter’s response times and the seeker’s response times for the Slow and Quick threads. We find that the peer-supporter’s response is, on average, 33 times slower than the seeker’s response in Short Repeated Seeker Interaction threads; it is only three times slower in the Quick counterparts. This contrast between the ratios and the corresponding retention likelihoods hint towards associations between response time of peer-supporters and their retention.

In order to better understand the dynamics between the two, we take a closer look at the first-time peer-supporters. We find that retention likelihood of peer-supporters is directly correlated with the response times in their first thread; a slow first response has a higher retention likelihood than a quick first response (Figure 7a). This indicates that first-time peer-supporters who are slower-to-act are more likely to return to the platform. These peer-supporters may be acting slowly due to multiple reasons. They might be carefully finding and selecting the threads to respond to; they might be taking more time to write; or they might be getting to know the platform interface. Disentangling these explanations is an important direction for future work.

### 7.4 Engaging in Mutual Discourse

We demonstrated that Mutual Discourse is an important pattern of engagement associated with higher seeker retention (Section 7.2) and peer-supporter retention (Section 7.3) likelihood. Next, we investigate when do threads

<sup>7</sup>Throughout the paper, we use Welch’s t-test for statistical testing unless stated otherwise.

## 8 Conclusion

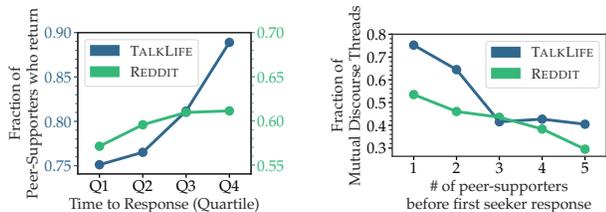
**Summary.** Online peer-to-peer support platforms facilitate mental health support but require users on the platforms to interact and engage. In this paper, we conducted a large-scale study of thread-level engagement patterns on two mental health support platforms, TALKLIFE & REDDIT. We operationalized four theory-motivated engagement indicators which were then synthesized into 11 distinct, interpretable patterns of engagement using a generative modeling approach. Our framework of engagement is multi-dimensional and models engagement jointly on the amount of the attention received by a thread and the interaction received by the thread. We then demonstrated how our framework of engagement patterns can be useful in evaluating the functioning of mental health platforms and for informing design decisions. We contrasted between TALKLIFE and REDDIT using their engagement pattern distributions which suggested that topically focused sub-communities, as found on REDDIT, may be important in making online support platforms more engaging. We found `Mutual Discourse` to be critical for seeker retention, particularly on TALKLIFE, facilitating which forms an important immediate future research direction.

**Risks and limitations.** Our study provides new insights on the patterns of engagement between seekers and peer-supporters and the associations of these patterns with their retention behavior. While these insights may have direct implications on the design of mental health support interventions, we note that our analysis is correlational and we cannot make any causal claims. Future work is needed to investigate the causal impact of engagement patterns and their effects on short-term and long-term individual health (Saha and Sharma 2020). Also, our framework of engagement does not account for content of posts, using which often involves ethical risks. This restricts the usage of certain popular interaction theories (e.g. (Sheizaf Rafaeli 1988)) in which the process of authoring of a post is dependent on the content of the previous posts in the thread.

Finally, we recommend that researchers and platform designers carefully consider the associated risks when considering interventions. For example, we found that `Mutual Discourse`, which has strong associations with both seeker and peer-supporter retention, is usually characterized by seekers replying back early to the thread. This may prompt platform designers to build, say an app which persuades support seekers to post early responses. However, such an intervention may have unintended consequences since we also found that the responses in `Mutual Discourse` often involve negative sentiment. Thus, persuading seekers to repeatedly report self-disclosures (De Choudhury and De 2014; Yang et al. 2019b) with negative sentiments may risk inducing negative effects on a potentially vulnerable population.

## Acknowledgments

We would like to thank TalkLife for providing us access to the licensed data and Sachin Pendse for help with initial data processing. We also thank Taisa Kushner, Mike Merrill, and



(a) The likelihood of peer- (b) Mutual Discourse more supporter retention increases if likely if the seeker responds they are slower-to-act. Response early, right after the first time is divided into quartiles in peer-supporter. the plot.

Figure 7

become a `Mutual Discourse` and what are the factors associated with it. We aim to gain insights on what type of seekers engage repeatedly with peer-supporters and what characteristics of seekers and their posts elicits mutual discourse. Specifically, we focus on threads which evolve into `Mutual Discourse` after a response from the seeker (the case of `Repeated Seeker Interaction`). We present our analysis and results next.

**Mutual Discourse more likely with seekers who write more and who express negative sentiments in their responses.** We analyze the first response of a seeker in a `Repeated Seeker Interaction` thread and compare it with the responses in `Mutual Discourse`. We extract the top phrases in both these type of responses separately using TopMine (El-Kishky et al. 2014). We observe that responses by seekers in `Mutual Discourse` contain phrases having more negative sentiment associated with them (e.g. *feel like shit*, *commit suicide*, *don't have friends*) relative to `Repeated Seeker Interaction`. We quantitatively analyze sentiment of the seeker responses using VADER (Hutto and Gilbert 2014). We find that the average sentiment in `Mutual Discourse` is more negative (TALKLIFE - 0.096 vs. 0.074; REDDIT - 0.078 vs. 0.064;  $p < 0.01$ ) and less positive (TALKLIFE - 0.152 vs. 0.222; REDDIT - 0.133 vs. 0.171;  $p < 0.001$ ) than in `Repeated Seeker Interaction`. This indicates that seekers who post negative sentiment in their responses are more likely to be involved in `Mutual Discourse`. We also find that the seeker responses in `Mutual Discourse`, on an average, contain more words (17.16 vs. 13.89;  $p < 0.001$ ).

**Mutual Discourse more likely when seekers respond early.** Next, we investigate when do seekers respond in a thread and if it correlates with the thread evolving into a `Mutual Discourse`. We find that threads where seekers respond right after the first peer-supporter are more likely to be `Mutual Discourse` (Figure 7b; 75.22% & 53.51% for TALKLIFE & REDDIT respectively). The likelihood tends to decrease if more number of peer-supporters reply in between. This can potentially be useful in designing support interventions in which the seeker is persuaded to respond early so that a `Mutual Discourse` is possible.

Koustuv Saha for their feedback on this work. Tim Althoff was funded in part by NSF grant IIS-1901386 and the Allen Institute for Artificial Intelligence.

## References

- Althoff, T., and Leskovec, J. 2015. Donor retention in online crowdfunding communities: A case study of donorschoose.org. In *WWW'15*, 34–44.
- Althoff, T.; Clark, K.; and Leskovec, J. 2016. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *TACL'16* 4:463–476.
- Andalibi, N., and Forte, A. 2018. Responding to sensitive disclosures on social media: A decision-making framework. *ACM TOCHI* 25(6):31.
- Andalibi, N.; Haimson, O. L.; Choudhury, M. D.; and Forte, A. 2018. Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. *ACM TOCHI* 25(5):28.
- Aragón, P.; Gómez, V.; and Kaltenbrunner, A. 2017. To thread or not to thread: The impact of conversation threading on online discussion. In *ICWSM'17*.
- Backstrom, L.; Kleinberg, J.; Lee, L.; and Danescu-Niculescu-Mizil, C. 2013. Characterizing and curating conversation threads: expansion, focus, volume, re-entry. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 13–22. ACM.
- Benton, A.; Mitchell, M.; and Hovy, D. 2017. Multitask learning for mental health conditions with limited social media data. In *EACL'17*, 152–162.
- Biyani, P.; Caragea, C.; Mitra, P.; and Yen, J. 2014. Identifying emotional and informational support in online health communities. In *COLING 2014*, 827–836.
- Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *JMLR* 3(Jan):993–1022.
- Bretz, R., and Schmidbauer, M. 1983. Media for interactive communication. *Sage Publications*.
- Collins, P. Y.; Patel, V.; Joestl, S. S.; March, D.; Insel, T. R.; Daar, A. S.; Bordin, I. A.; Costello, E. J.; Durkin, M.; Fairburn, C.; et al. 2011. Grand challenges in global mental health. *Nature* 475(7354):27.
- Coppersmith, G.; Harman, C.; and Dredze, M. 2014. Measuring post traumatic stress disorder in twitter. In *ICWSM'14*.
- Davidson, L.; Chinman, M.; Kloos, B.; Weingarten, R.; Stayner, D.; and Tebes, J. K. 1999. Peer support among individuals with severe mental illness: A review of the evidence. *Clinical psychology: Science and practice* 6(2):165–187.
- De Choudhury, M., and De, S. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *ICWSM'14*.
- De Choudhury, M., and Kiciman, E. 2017. The language of social support in social media and its effect on suicidal ideation risk. In *ICWSM'17*.
- De Choudhury, M.; Gamon, M.; Counts, S.; and Horvitz, E. 2013. Predicting depression via social media. In *ICWSM'13*.
- De Choudhury, M.; Counts, S.; Horvitz, E. J.; and Hoff, A. 2014. Characterizing and predicting postpartum depression from shared facebook data. In *CSCW'14*, 626–638. ACM.
- De Choudhury, M.; Kiciman, E.; Dredze, M.; Coppersmith, G.; and Kumar, M. 2016. Discovering shifts to suicidal ideation from mental health content in social media. In *CHI'16*, 2098–2110. ACM.
- Diao, Q.; Jiang, J.; Zhu, F.; and Lim, E.-P. 2012. Finding bursty topics from microblogs. In *ACL'12*, 536–544.
- El-Kishky, A.; Song, Y.; Wang, C.; Voss, C. R.; and Han, J. 2014. Scalable topical phrase mining from text corpora. *Proceedings of the VLDB Endowment* 8(3):305–316.
- Ernala, S. K.; Labetoulle, T.; Bane, F.; Birnbaum, M. L.; Rizvi, A. F.; Kane, J. M.; and De Choudhury, M. 2018. Characterizing audience engagement and assessing its impact on social media disclosures of mental illnesses. In *ICWSM'18*.
- Gaur, M.; Alambo, A.; Sain, J. P.; Kursuncu, U.; Thirunarayan, K.; Kavuluru, R.; Sheth, A.; Welton, R.; and Pathak, J. 2019. Knowledge-aware assessment of severity of suicide risk for early intervention. In *WWW'19*, 514–525. ACM.
- Hutto, C. J., and Gilbert, E. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM'14*.
- Kaplan, B. H.; Cassel, J. C.; and Gore, S. 1977. Social support and health. *Medical care* 15(5):47–58.
- Kumar, R.; Mahdian, M.; and McGlohon, M. 2010. Dynamics of conversations. In *SIGKDD'10*, 553–562. ACM.
- Lin, H.; Jia, J.; Guo, Q.; Xue, Y.; Li, Q.; Huang, J.; Cai, L.; and Feng, L. 2014. User-level psychological stress detection from social media using deep neural network. In *Proceedings of the 22nd ACM international conference on Multimedia*, 507–516. ACM.
- Lumbreras, A.; Jouve, B.; Velcin, J.; and Guégan, M. 2017. Role detection in online forums based on growth models for trees. *Social Network Analysis and Mining* 7(1):49.
- Mak, S.; Williams, R.; and Mackness, J. 2010. Blogs and forums as communication and learning tools in a mooc. In *Proceedings of the 7th International Conference on Networked Learning 2010*, 275–285. University of Lancaster.
- Mayshak, R.; Sharman, S. J.; Zinkiewicz, L.; and Hayley, A. 2017. The influence of empathy and self-presentation on engagement with social networking website posts. *Computers in Human Behavior* 71:362–377.
- O'Brien, H. L., and Toms, E. G. 2008. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology* 59(6):938–955.
- Pérez-Rosas, V.; Wu, X.; Resnicow, K.; and Mihalcea, R. 2019. What makes a good counselor? learning to distinguish between high-quality and low-quality counseling conversations. In *ACL'19*, 926–935.

- Peruani, F.; Choudhury, M.; Mukherjee, A.; and Ganguly, N. 2007. Emergence of a non-scaling degree distribution in bipartite networks: a numerical and analytical study. *EPL (Europhysics Letters)* 79(2):28001.
- Pruksachatkun, Y.; Pendse, S. R.; and Sharma, A. 2019. Moments of change: Analyzing peer-based cognitive support in online mental health forums. In *CHI'19*, 64. ACM.
- Rathod, S.; Pinninti, N.; Irfan, M.; Gorczynski, P.; Rathod, P.; Gega, L.; and Naeem, F. 2017. Mental health service provision in low-and middle-income countries. *Health services insights* 10:1178632917694350.
- Reece, A. G., and Danforth, C. M. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Science* 6(1):15.
- Ridley, C. A., and Avery, A. W. 1979. Social network influence on the dyadic relationship. In *Social exchange in developing relationships*. Elsevier. 223–246.
- Rosengren, K. E. 1999. *Communication: an introduction*. Sage.
- Saha, K., and Sharma, A. 2020. Causal factors of effective psychosocial outcomes in online mental health communities. In *ICWSM'20*.
- Sharma, E., and De Choudhury, M. 2018. Mental health support and its relationship to linguistic accommodation in online communities. In *CHI'18*, 641. ACM.
- Sheizaf Rafaeli, S. 1988. Interactivity: From new media to communication. *Beverly Hill in Hawkins, Pingree, Wieman Edition*.
- Smith, W.; Ploderer, B.; Wadley, G.; Webber, S.; and Borland, R. 2017. Trajectories of engagement and disengagement with a story-based smoking cessation app. In *CHI'17*, 3045–3056. ACM.
- Taylor, L. D. 2011. Avatars and emotional engagement in asynchronous online communication. *Cyberpsychology, Behavior, and Social Networking* 14(4):207–212.
- Tsugawa, S.; Kikuchi, Y.; Kishino, F.; Nakajima, K.; Itoh, Y.; and Ohsaki, H. 2015. Recognizing depression from twitter activity. In *CHI'15*, 3187–3196. ACM.
- Van Uden-Kraan, C. F.; Drossaert, C. H.; Taal, E.; Smit, W. M.; Moens, H. J. B.; and Van de Laar, M. A. 2011. Determinants of engagement in face-to-face and online patient support groups. *JMIR* 13(4):e106.
- Wang, X., and McCallum, A. 2006. Topics over time: a non-markov continuous-time model of topical trends. In *SIGKDD'06*, 424–433. ACM.
- Wang, C.; Ye, M.; and Huberman, B. A. 2012. From user comments to on-line conversations. In *SIGKDD'12*, 244–252. ACM.
- White, M., and Dorman, S. M. 2001. Receiving social support online: implications for health education. *Health education research* 16(6):693–707.
- Williams, F.; Rice, R. E.; and Rogers, E. M. 1988. *Research methods and the new media*. Simon and Schuster.
- Yang, D.; Kraut, R. E.; Smith, T.; Mayfield, E.; and Jurafsky, D. 2019a. Seekers, providers, welcomers, and storytellers: Modeling social roles in online health communities. In *CHI'19*, 344. ACM.
- Yang, D.; Yao, Z.; Seering, J.; and Kraut, R. 2019b. The channel matters: Self-disclosure, reciprocity and social support in online cancer support groups. In *CHI'19*, 31. ACM.
- Yin, J., and Wang, J. 2014. A dirichlet multinomial mixture model-based approach for short text clustering. In *SIGKDD'14*, 233–242. ACM.
- Yu, J.; Hu, Y.; Yu, M.; and Di, Z. 2010. Analyzing netizens view and reply behaviors on the forum. *Physica A: Statistical Mechanics and its Applications* 389(16):3267–3273.