PLOS ONE

Critical behavioral traits foster peer engagement in Online Mental Health Communities

Aseem Srivastava[1]*, Tanya Gupta[1], Alison Cerezo[2,3], Sarah Peregrine (Grin) Lord[2,4], Md Shad Akhtar[1], Tanmoy Chakraborty[5,6]

1 Department of Computer Science and Engineering, Indraprastha Institute of Information Technology Delhi, New Delhi, India, 2 Empathy Rocks, Inc. dba mpathic.ai, Bellevue, WA, United States of America, 3 Department of Counseling, Clinical and School Psychology, University of California Santa Barbara, Santa Barbara, California, United States of America, 4 Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States of America, 5 Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India, 6 Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, New Delhi, India

* aseems@iiitd.ac.in

Abstract

Online Mental Health Communities (OMHCs), such as Reddit, have witnessed a surge in popularity as go-to platforms for seeking information and support in managing mental health needs. Platforms like Reddit offer immediate interactions with peers, granting users a vital space for seeking mental health assistance. However, the largely unregulated nature of these platforms introduces intricate challenges for both users and society at large. This study explores the factors that drive peer engagement within counseling threads, aiming to enhance our understanding of this critical phenomenon. We introduce BeCOPE, a novel behavior encoded Peer counseling dataset comprising over 10, 118 posts and 58, 279 comments sourced from 21 mental health-specific subreddits. The dataset is annotated using three major fine-grained behavior labels: (a) intent, (b) criticism, and (c) readability, along with the emotion labels. Our analysis indicates the prominence of "self-criticism" as the most prevalent form of criticism expressed by help-seekers, accounting for a significant 43% of interactions. Intriguingly, we observe that individuals who explicitly express their need for help are 18.01% more likely to receive assistance compared to those who present "surveys" or engage in "rants." Furthermore, we highlight the pivotal role of well-articulated problem descriptions, showing that superior readability effectively doubles the likelihood of receiving the sought-after support. Our study emphasizes the essential role of OMHCs in offering personalized guidance and unveils behavior-driven engagement patterns.

1 Introduction

The prevalence of mental health distress has risen sharply in the last several years. A recent report reveals that one in six individuals suffers from mental health-related challenges

---

https://www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care. At the same time, there is a severe shortage of mental health providers to facilitate adequate support to those in need https://www.newamericaneconomy.org/press-release/new-study-shows-60-percent-of-u-s-counties-without-a-single-psychiatrist/ [1, 2]. As a result of these growing challenges, we specifically examined the patterns and factors that drive individuals to engage with peer-to-peer mental health threads, focusing on the impact of behavioral, emotional, textual, and topical signals during peer-to-peer interactions.

To this end, we develop the BeCOPE (BEhavior enCOded PEer Counseling) dataset, composed of peer-to-peer mental health conversational interactions across 10, 118 posts and 58, 279 comments from 21 mental health-specific subreddits. We inspect the level of engagement on Reddit for three different OMHC categories—(a) interactive, (b) non-interactive, and (c) isolated—based on the pattern of interaction between users and the original help-seeker (see Fig 1). Analyzing the critical factors in each engagement category, we comprehend factors and patterns that lead to constructive versus detrimental peer-to-peer mental health interactions. Understanding peer-to-peer interactions on OMHCs is key to the ethical and safe monitoring of these communities, including the moderation of safe interactions and the sharing of accurate mental health information. We explore the following research questions:

RQ1. When examining peer-to-peer OMHC interactions, how do intent (i.e., help-seeking), readability, and criticism impact peer willingness to engage with the original post (e.g., validation, advice-giving)?

RQ2. How does the expression of emotions in posts impact user engagement in the OMHC platforms?
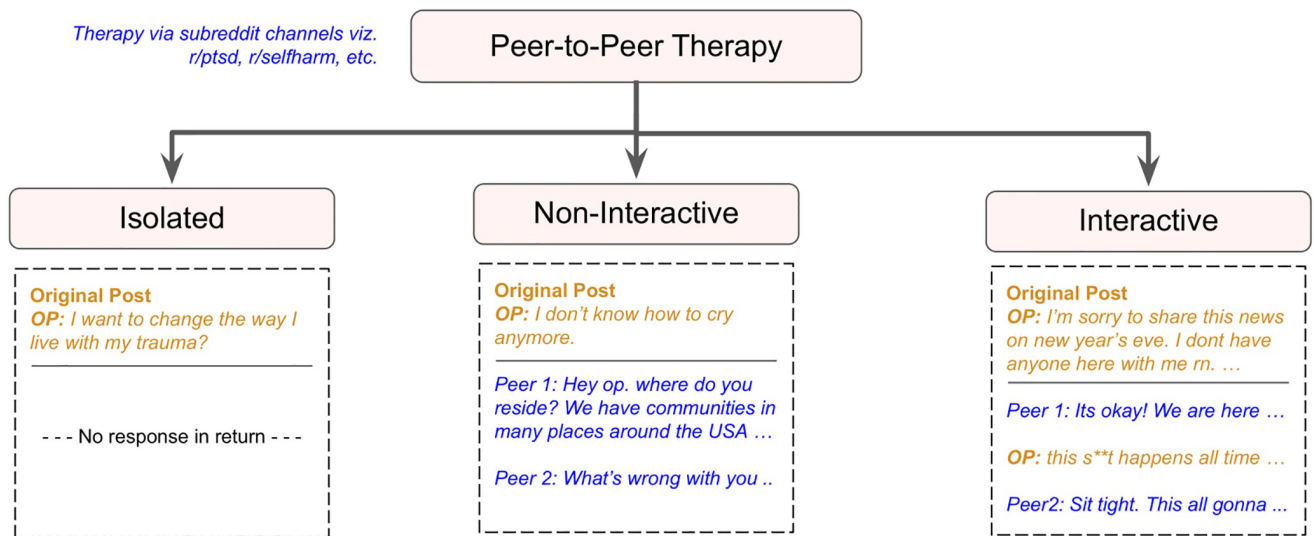
Fig 1. Taxonomy of counseling methods along with examples. Here, OP (original poster) is a common Internet terminology for the person who creates posts on peer-to-peer platforms. In peer-to-peer therapy, we inspect the level of engagement in three different categories based on the abundance of interaction with the help-seeker—**(a) interactive:** if there are back-and-forth conversations between the OP and peers, **(b) non-interactive:** if the post engages peers, but the OP does not reply to peers, and **(c) isolated:** if the post does not have any comment, but one-to-one therapy involves the continuous exchange of dialogues between therapist and client (help-seeker).

https://doi.org/10.1371/journal.pone.0316906.g001

Reddit is a popular OMHC platform that has steadily emerged as a platform for seeking help concerning a spectrum of mental challenges with specific posts devoted to disorders such as depression, attention-deficit/ hyperactivity disorder (ADHD, sometimes ADD), bipolar disorder and alcohol and substance use [3–5]. Typically, users (i.e., support-seekers) create original posts to discuss their mental health issues, describing their symptoms and the contexts of their specific situations, like job loss or a recent divorce. The support-seekers, in turn, receive replies from peers (other users on the platform) with advice, recommendations for symptom management, and general support. This process allows support-seekers to share and ask for help for their mental health challenges in a cost-effective, convenient, and anonymous manner that typically results in immediate support. A recent study [6] analyzed patterns of posts on two popular OMHC platforms, Talklife and Reddit, by leveraging natural language processing for communication models in human-computer interaction and communication theory, operationalizing a set of four engagement indicators based on attention and interaction. The authors found that the back-and-forth peer platform communication effectively contributes to early support. A similar study [7] examined the change in sentiment to analyze peer-to-peer counseling settings to read whether a counseling thread or a post on the platform is correlated with a moment of cognitive change. It turned out that behavioral signals such as sentiment, affect, and topics associated with language are decisive toward effective counseling. On the same track, another study discussed the temporal engagement on social media correlating with patient disclosure [8]. The authors developed an autoregressive time series computational model that assesses engagement patterns and subsequently forecasts alteration in the intimacy of disclosures. They found that attributes of audience engagement, like emotional support, personal behavior, and self-disclosure, strongly predict patterns in future counseling behavior.

Previous studies on the analysis of peer-to-peer mental health interactions identified threads that fall into affective [5], content-based [9, 10], and supportive [11] categories, thus demonstrating reliability for the functioning of peer-to-peer mental health platforms. However, little is known about how these categories of peer-to-peer mental health interactions are associated with constructive and/or detrimental outcomes. Understanding the characteristics of such OMHC users [12–15] and given the widespread use of OMHC platforms, specific patterns and factors that drive engagement in peer-to-peer mental health interaction must be identified [6, 16, 17]. In doing so, social media platforms should be better able to monitor and intervene for the benefit of their users in distress [18–20].

In summary, these studies examine engagement patterns based on interaction indicators with and without the original poster's (OP) presence, attempting to model these indicators for better understanding and the impact of peer-based cognitive support on users who initially seek support, showing how such interactions can lead to positive outcomes. Other works focus on patient disclosure for engagement patterns based on the nature of information shared by users. These studies provide valuable insights into engagement criteria and user behavior but also leave a research gap in addressing the specific characteristics and patterns of engagement we investigate. We take collective inspiration from these studies to define engagement categories and identify research gaps. Our study introduces novel research questions that have not been explored before, focusing on the specific factors driving engagement and the characteristics of users in mental health-specific subreddits.

2 Methods

2.1 Data collection

To study latent signals in peer-to-peer mental health interactions, we develop BeCOPE by curating posts from 21 subreddits. Reddit is organized into spaces called subreddits, where

each subreddit is specific to a certain discussion topic. To analyze behaviors on peer-to-peer mental health platforms, we scraped, processed, and annotated subreddit data to develop the dataset.

2.1.1 Selection of subreddits. We operationalized an initial screening of diverse subreddits by considering the posting activity during the COVID-affected year 2020 across common mental health discussion areas. Our experts identified the shortlisted subreddits with maximal coverage and diversity. The selection of 21 subreddits, as shown in Table 1, is to ensure a broad yet manageable dataset that provides comprehensive insights while maintaining analytical feasibility. The selection process aimed to capture a diverse range of interactions across various communities without overwhelming the analysis process. For each shown subreddit, we curated 500 posts and their comments from January 2020 to December 2020. This is worth noting, considering the activity on these subreddits, we did not encounter any case where posts fall short of 500. Further, we performed a sanity check to ensure that conversations were acceptable (e.g., noise-free, written in English). In total, we collected 10, 118 posts and 58, 279 comments along with their metadata, such as author information, score (upvotes), time of creation, and the number of comments.

Step 1: Categorization of interactions by the level of peer engagement. Depending on the comments on a post, we classified the collected conversations into one of the three engagement categories: (i) interactive, (ii) non-interactive, or (iii) isolated. If an original post involved back-and-forth comments from the original user and peers, the conversation was deemed

Table 1. Statistics of the BeCOPE dataset. We collected a total of $\sim 10K$ posts and $\sim 50K$ comments. We annotated all the posts using three core labels—(i) intent, (ii) criticism, and (iii) readability (Clear: Excellent, Good, and Average; Non-clear: Mediocre and Poor). IAA (κ) represents the inter-annotator agreement using Cohen's kappa score.

Subreddits	Posts	Comments	Intent				Criticism				Readability	
			Help seeking	Rant	Survey	Chitchat	Self crit	Other w/ Rsn	Other w/o Rsn	No criticism	Clear	Non-clear
r/Anxiety	469	1773	252	129	62	26	278	48	7	136	467	2
r/ptsd	494	1567	221	144	64	65	180	135	1	178	494	0
r/suicideWatch	403	2545	90	246	17	50	231	34	10	128	378	25
r/addiction	487	3581	217	148	43	79	246	67	6	168	466	21
r/ADHD	423	3856	169	104	78	72	139	31	9	247	418	5
r/alcoholicsanonymous	498	6021	181	107	47	163	155	58	5	280	490	8
r/Anger	464	2620	233	184	31	16	245	140	16	63	462	2
r/BPD	519	2744	180	185	113	41	234	99	4	182	518	1
r/depression	547	1951	83	363	26	75	243	91	18	195	546	1
r/domesticviolence	425	2847	254	94	25	52	34	277	1	113	421	4
r/eating_disorders	568	2021	256	209	51	52	346	43	1	178	567	1
r/getting_over_it	476	2551	230	163	35	48	258	72	2	144	473	3
r/mentalillness	484	1895	208	155	52	69	209	99	2	174	480	4
r/OpiatesRecovery	493	6112	215	116	62	100	185	28	3	277	493	0
r/rapecounseling	481	2390	288	142	26	25	125	269	1	86	481	0
r/sad	486	2258	44	287	27	128	115	71	8	292	485	1
r/selfharm	467	1928	136	232	52	47	243	39	0	185	465	2
r/selfhelp	419	2001	177	60	28	154	163	37	0	219	390	29
r/socialanxiety	461	2798	167	128	64	102	201	58	0	202	428	33
r/OCD	424	2528	159	117	63	85	209	29	3	183	424	0
r/helpmecope	473	2121	277	127	17	52	170	160	2	141	471	2
Total	**9961**	**58108**	**4037**	**3440**	**983**	**1501**	**4209**	**1885**	**99**	**3771**	**9817**	**144**
IAA (κ)	**-**	**-**	**0.963**				**0.888**				**0.861**	

Table 2. Example of posts and their corresponding labels in BeCOPE. Intent:*Help-seeking*, *Rant*, *Chit-chat*, and *Survey*; Criticism:*Self-criticism* (SC), *Criticism with reasoning* (CR), *Criticism with no reasoning* (CNR), and *No-criticism* (NC); Readability:*Excellent* (5), *Good* (4); *Average* (3), *Mediocre* (2), and *Poor* (1); Emotion:*Admiration, Amusement, Anger, Annoyance, Approval, Caring, Confusion, Curiosity, Desire, Disappointment, Disapproval, Disgust, Embarrassment, Excitement, Fear, Gratitude, Grief, Joy, Love, Nervousness, Optimism, Pride, Realization, Relief, Remorse, Sadness, Surprise*, and *Neutral*; Engagement:*Interactive, Non-interactive*, and *Isolated*.

| Post	Intent	Criticism	Emotion		Read
			Primary	**Secondary**	
Hello all. I unfortunately used again today, despite going to my first NA meeting last night. My loneliness, is a trigger. I spent the whole day by myself, and decided it wouldn't hurt to light up one more time. I have a job interview, next Monday. I don't want to lose my life. I am usually a proud person, and solve my problems on my own. But, I know I need to reach out to someone before I spiral out of control. I honestly see myself giving up everything for it. Please, any advice and tips would help me in this low moment. I have no one in my life currently. Thank you. [**Interactive**]	*Help-seeking*	SC	*Sadness*	*Gratitude*	*Excellent*
I recently found out my boyfriend has been crushing or opening then snorting most of his prescription medications. We had been arguing non-stop over his marijuana addiction, but this brings it to the next level. Not just his ADHD meds, but his depression and anxiety meds too. He claims it makes them more effective. I tried negotiating with him over letting me dispense his ADHD meds to him one week at a time, and he goes on tirades about loss of control/lack of trust/I'm not his mother/the government and his doctors can't control him etc. I already am exhausted from our constant fighting over marijuana abuse. He says it's not my problem because it's his body and his decision. He has ADHD, depression, and anxiety concurrent with substance abuse disorder (marijuana and alcohol). He has been hospitalized and gone to rehabilitation multiple times. Now he's adding RX abuse to his problems. I told him I was considering reporting him to his doctors, and he told me to keep my mouth shut and let him make his own decisions. I think his illness keeps him from making logical decisions about this, as he is constantly seeking ways to "not be bored", "kill the anxiety", or "not feel anything". TLDR: SO says his decision to abuse prescription medications is not my problem, despite his diagnosed substance abuse disorders.[**Interactive**]	*Rant*	CR	*Sadness*	*Disappointment*	*Average*
I live in Canada. I started the application process for medical assistance in dying. It was nice to finally receive an intelligible response to "I really want to die." as oppose to the usual parroted phrases and hollow cheerleading. Not sure that I will go through with it tbh, but I hope that I do. Getting two witness signatures seems like the only real obstacle here. [**Isolated**]	*Chitchat*	CNR	*Optimism*	*Joy*	*Excellent*
Who's an alcoholic stay at home wife/mom? Just super interesting to me. What are your days typically like? Do you hide your consumption or own it? [**Non-Interactive**]	*Survey*	NC	*Excitement*	*Neutral*	*Excellent*
Every day I read to you from Daily Reflections. This book is published by Alcoholics Anonymous, and is an important resource for members of this LIFE-saving association. Many recovering alcoholics use this literature to start their day. AA is a fellowship of men and women who share their strengths and hopes with each other for the common purpose of helping the alcoholic who still suffers. It is the faith and the love of the LIFE which allows the cure or rather ashes the progression of the disease and helps us to leave the hell of alcoholism one day at the time. Addiction is very powerful and sneaky. Although these writings are primarily intended for alcoholics, their families and friends, many people who feel they have no contact with alcoholism greatly appreciate the wisdom that emanates from them. We hope you will gain the freedom we know! MERCI! Thank you for your support by subscribing to our YouTube channel. <youtube channel> #recovery #alcoholism #alcoholics anonymous love #LIFE #January [**Isolated**]	*Chitchat*	NC	*Neutral*	*Neutral*	*Poor*

https://doi.org/10.1371/journal.pone.0316906.t002

"interactive" (see Section 2.2 and Table 2) for an example). If an original post had zero comments, the conversation was deemed "isolated." Finally, if an original post received more than one comment from peers, but the original user did not acknowledge or reply to peers' comments, the conversation was deemed "non-interactive".

The process of determining factors that impact peer engagement was a collaborative effort made by domain experts. The three broad scenarios that we discussed were carefully analyzed to identify patterns and factors that could be associated with higher levels of engagement. By considering all possible outcomes, we were able to narrow down the analysis to a focused set of factors that significantly correlated with user engagement.

Step 2: Annotation of posts by behavioral and emotional labels. The first step in the annotation process was the curation of Reddit posts on mental health topics by categorizing them based on (i) intent, (ii) criticism, (iii) readability, and (iv) emotion labels. We manually

annotated ∼5K posts and subsequently learned respective classifiers to obtain pseudo-labels for another ∼5K posts by finetuning BERT [21]. We present Fig 2(a) shows the performance of the pseudo-labeling process. Next, a sanity check of the annotated dataset was performed to ensure the reliability of the annotations. Finally, we used the resultant dataset of ∼10 posts for our analyses. Detailed statistics of the annotated BeCOPE dataset, after pseudo labeling, are presented in Table 1 and Fig 2(b). Next, we discuss the annotation details.
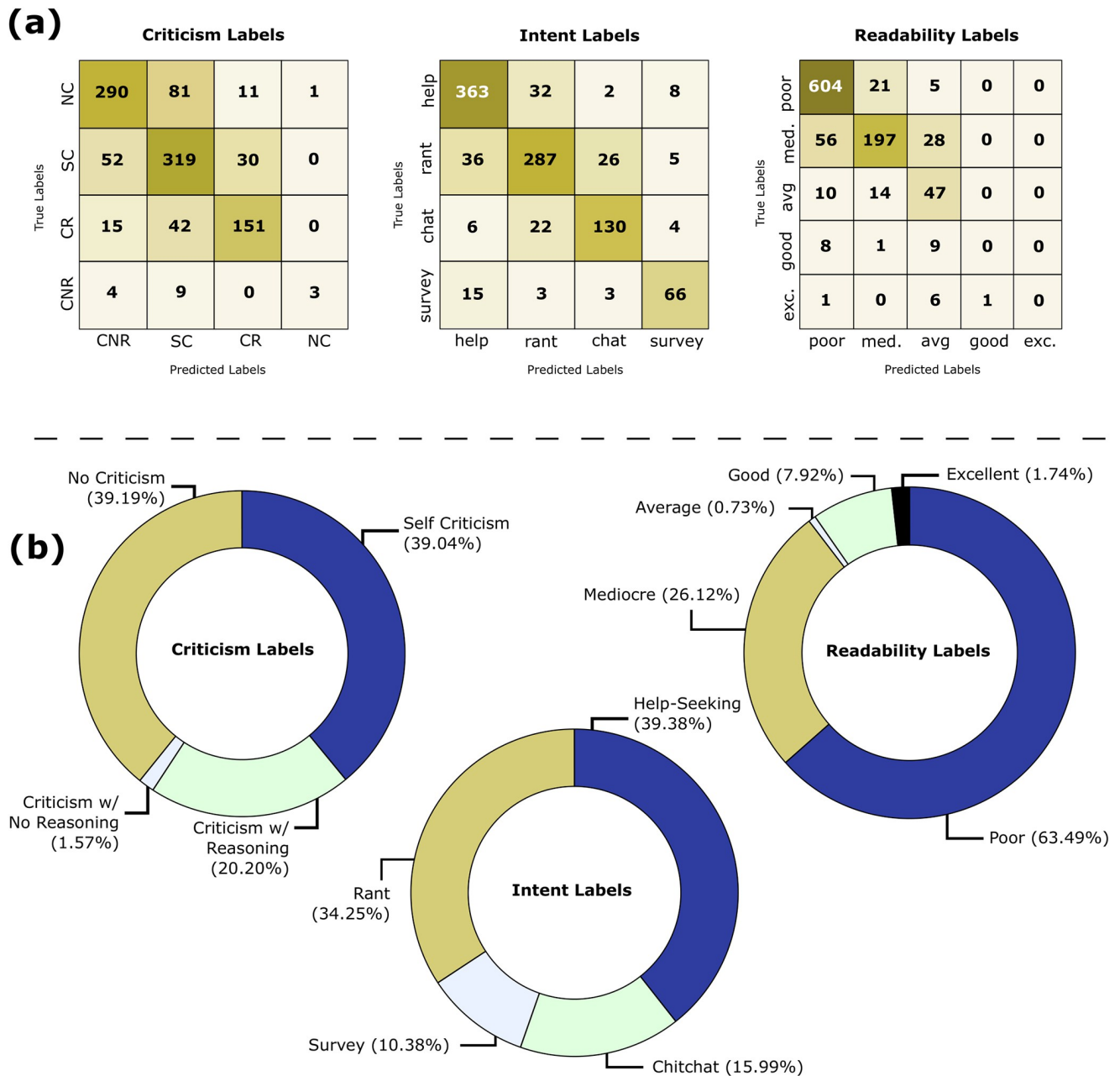
Fig 2. (**a**) Confusion matrix to represent the performance of pseudo labeling of criticism, intent, and readability labels. We exploit BERT to fine-tune on ∼5K manually annotated posts to predict criticism, intent, and readability on the remaining posts. (**b**) Distribution of behavioral signals (criticism and intent) along with readability in the complete BeCOPE dataset.

2.2 Data annotation

Peer-to-peer counseling conversations are open-ended, where users express their diverse and different perspectives. We observe that users exhibit a variety of intents while discussing mental health issues *viz.* rating, seeking help survey, or doing general chit-chat. On the other hand, some alleged criticism for their issues. Moreover, another aspect of peer-to-peer counseling is to understand whether users convey their needs clearly and crisply. A well-written post may have attracted more and perhaps better responses than a poorly-written post. Therefore, the readability of the post is another factor that directly affects the chances of receiving help on a mental health post from peers. At the same time, studying the expressed emotions is highly imperative. We hypothesize that all the above-discussed factors—the knowledge of intent, the presence of criticism, the study of emotions, and the readability of the posts, are crucial in understanding the need of the help-seeker and accordingly providing appropriate assistance.

Considering the literature and observations, we designed a set of guidelines to annotate the curated Reddit posts. A detailed discussion of the guidelines considering the four factors and their definition is presented below.

2.3 Intent

Intent defines the purpose of the original poster (OP) in the post. We divide the posts into four categories based on the user's needs: help-seeking, rant, survey, and chitchat.

- Help-Seeking: Original posters explain mental health issues and expect peers to provide helpful suggestions to improve their condition.

- Rant: Original posters share their (strong) views on mental health issues without expecting help from peers.

- Survey: Original posters share mental health issues and ask peers to share their experiences. *Survey* differs from *help-seeking* as survey-labeled posts ask for a generic point of view on related mental health issues rather than individual-centric assistance.

- Chitchat: The Chitchat label is used for filler posts that are not directly related to mental health issues. Such posts include well wishes, general guidelines, occasional greetings, etc.

2.4 Criticism

Original posters often criticize the situation caused due to their or others' mental health issues. Sometimes, criticism is on their own; other times, it is on others. Hence, it is important to study if showing criticism could be a cause to receive better help. In other words, do peers prefer helping others who use criticizing language in posting their mental state? To understand this, we define four criticism labels: *no-criticism, self-criticism, others' criticism with reason* and *without reason*.

- Self-Criticism (SC). We use this label for posts where original posters criticize themselves for their mental health issues.

- Criticism on Others with Reason (CR). We use this label for posts where original posters criticize others for their mental health issues. Also, they provide some reasons (justification) to support their criticism.

- Criticism on Others with No Reason (CNR). This label differs from CR as the criticism is not backed by reasoning.

- No-Criticism (NC). We use this label for posts where there is no criticism.

2.5 Readability

Readability is essential in imparting most of the information via textual communication in all professional domains. Earlier works utilized the readability criteria to decide the impact of a mental health post using statistical properties of posts such as the length of the post. However, we argue that a shorter post could also be interpreted as poorly readable. Therefore, in this work, we define the readability score based on the clarity of the text and the amount of effort one needs to put into comprehending the post. We employed linguistic experts to analyze the general writing nuances of the original posters (OPs) and found that experts could distinctly differentiate between clear and non-clear posts. We observe that lengthier sentences pose a degree of uneasiness in readers besides the use of SMS slang and abbreviations. Based on the observations, we define five readability levels for a post—*excellent*, *good*, *average*, *mediocre*, and *poor*.

2.6 Emotion

Emotion labeling is the practice of cultivating empathetic knowledge in conversations. We employ a set of 28 emotion classes –*admiration, amusement, anger, annoyance, approval, caring, confusion, curiosity, desire, disappointment, disapproval, disgust, embarrassment, excitement, fear, gratitude, grief, joy, love, nervousness, optimism, pride, realization, relief, remorse, sadness, surprise*, and *neutral*– for our Reddit posts. Moreover, we observed that many posts conveyed multiple emotions in a single post; hence, we assigned two emotions for each post, i.e., the *primary emotion* and the *secondary emotion*.

2.7 Annotation quality

To ensure the rigor of our annotations, we implemented an iterative process. This was designed to maintain high standards of consistency across complete data. Initially, we defined clear annotation guidelines for each category, drawing from established literature and expert consultations. These guidelines were then used to train our annotators, who were linguistic experts as well as mental health experts who were native English speakers. Our training sessions included detailed guidelines and examples to help annotators accurately identify posts based on each annotation category. We employed a series of iterative revisions and feedback loops to refine our annotations. Annotators worked asynchronously in a sequential manner, and their feedback was recorded and reviewed after every 200 posts. This iterative process allowed us to identify and address any discrepancies or ambiguities in the annotation criteria. To quantify the reliability of our annotations, we calculated the Cohen's Kappa score for each category. After four (4) rounds of revisions and training, we achieved substantial agreement scores: a Cohen's Kappa of 0.963 for intent, 0.888 for readability, and 0.861 for emotion labels. These scores indicate a high level of agreement among annotators, reflecting the rigor and reliability of our annotation process.

2.7.1 Rationale behind the selection of labels.   The categories of intent, criticism, readability, and emotion labels are chosen based on a deliberate review of literature and consultation with experts in linguistics and mental health. These categories are critical for understanding the dynamics of online mental health discussions and their impact on user engagement. Precisely, intent captures the purpose behind posts and is categorized as help-seeking, ranting, surveying, and chitchat. This category helps differentiate between different types of user interactions and their underlying motivations. Criticism identifies negative interactions either directed at oneself or others, which is essential for understanding the tone and potential impact of discussions. We further categorized criticism as self-criticism, criticizing others with reasoning, criticizing others without reasoning, and no criticism, to capture the

nuances of such interactions. Readability assesses the ease of understanding the posts, which is crucial for engagement, as more readable posts are likely to attract more interaction. We categorized readability on a Likert scale of excellent, good, average, mediocre, and poor. This categorization helps in analyzing how the complexity of language influences user engagement. Emotion labels capture the affective nuances of the content, providing insights into the underlying emotions in the discussions. This is important for understanding the overall sentiment in the engagement categories.
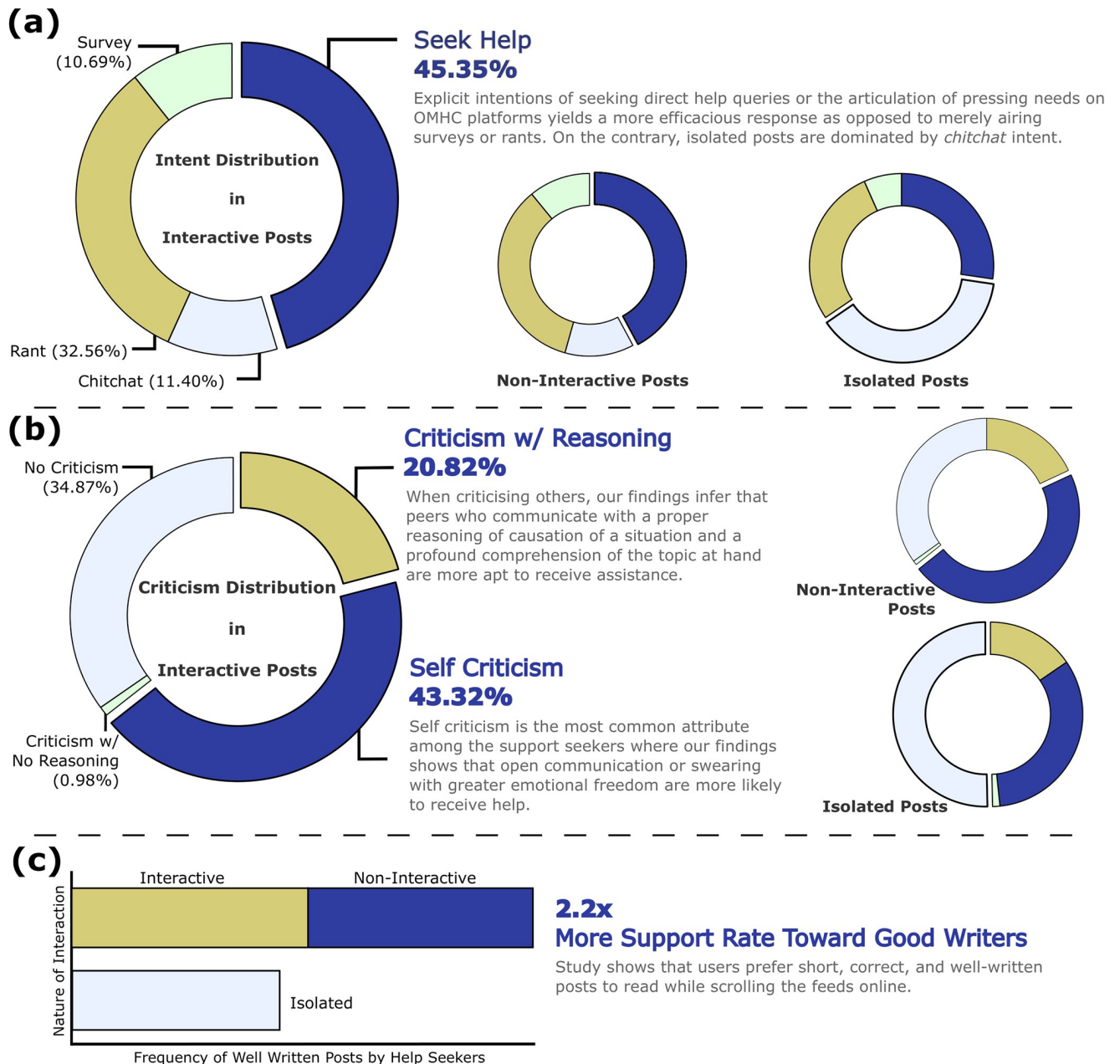
2.8 Ethical consideration

Considering the sensitivity of research in mental health, this paper does not include any personal, identifiable information of any OMHC user. We collected data solely based on the most relevant mental health subreddits and did not include any bias in the choice of particular subreddit channels. Finally, we conducted all experiments without compromising the anonymity of online users in BeCOPE.

3 Results and analysis

3.1 RQ1: When examining peer-to-peer OMHC interactions, how do intent (i.e., help-seeking), readability, and criticism impact peer willingness to engage with the original post (e.g., validation, advice giving)?

3.1.1 Intent. We observe that help-seekers on OMHC platforms are 18.01% more likely to receive help when they explicitly convey their pressing needs through queries, as opposed to when they make statements about their experiences. When an original post contains a help-seeking approach, it increases peer engagement. Specifically, 45.35% of interactive posts, 42.16% of non-interactive, and 27.34% of isolated posts are help-seeking in nature, indicating that peers who explicitly ask for help for their mental issues experience greater peer engagement. We also observe that when an original post is constructed as a "rant" (a long statement of the problem with no explicit ask for help/advice), it receives less peer engagement. The number of isolated posts labeled with the rant intent (38.11%) exceeds non-interactive posts (34.73%) and interactive posts (32.56%) by a margin of 3.38% and 5.55%, respectively. Further, posts with rant intent receive the least interaction compared to other intent labels across all engagement categories, showing that the survey posts do not elicit peers' attention toward assistance. Our analysis sheds light on RQ1 by indicating the conveyance of explicit intentions through queries or the articulation of pressing needs on the OMHC platforms yields a more efficacious response. We present the distribution of intents across three engagement categories in Fig 3(a). The four annotated intent labels receive a significant agreement score with a confidence ≥95% on the *p*-values of help-seeking (0.022), rant (0.046), chitchat (0.016), and survey (0.028). Furthermore, Section 2.2 presents fine-grained details of intent labels and their annotation.

3.1.2 Criticism. We observe that isolated posts have maximum *no-criticism* (NC) labels (50.34%) as compared to non-interactive (34.92%) and interactive (34.87%) posts. Fig 3(b) shows the distribution of *criticism* labels across all engagement categories. Conversely, individuals who can obtain support from their peers on OMHCs are frequently found to engage in criticising themselves and others. We bifurcate the criticism of others into two indicative categories—*criticism with reasoning* (CR) (i.e., a logical presentation of one's experience), and *criticism with no-reasoning* (CNR). Out of all three engagement categories, interactive engagement carries the maximum CR label, 2.75% and 5.39% more than non-interactive and isolated engagement categories, respectively. This trend directly draws attention to the fact that proper

(a)

Survey (10.69%)

Seek Help
45.35%

Explicit intentions of seeking direct help queries or the articulation of pressing needs on OMHC platforms yields a more efficacious response as opposed to merely airing surveys or rants. On the contrary, isolated posts are dominated by *chitchat* intent.

Intent Distribution in Interactive Posts

Rant (32.56%)

Chitchat (11.40%)

Non-Interactive Posts

Isolated Posts

(b)

No Criticism (34.87%)

Criticism w/ Reasoning
20.82%

When criticising others, our findings infer that peers who communicate with a proper reasoning of causation of a situation and a profound comprehension of the topic at hand are more apt to receive assistance.

Criticism Distribution in Interactive Posts

Self Criticism
43.32%

Self criticism is the most common attribute among the support seekers where our findings shows that open communication or swearing with greater emotional freedom are more likely to receive help.

Criticism w/ No Reasoning (0.98%)

Non-Interactive Posts

Isolated Posts

(c)

Interactive   Non-Interactive

Nature of Interaction

Isolated

Frequency of Well Written Posts by Help Seekers

2.2x
More Support Rate Toward Good Writers

Study shows that users prefer short, correct, and well-written posts to read while scrolling the feeds online.

Fig 3. Distribution of behavioral signals and readability in BeCOPE across all engagement categories. (a) The intent distribution indicates that a majority (45.35%) of posts show explicit intentions (seek-help) through queries or the articulation of pressing needs on OMHC platforms, yielding productive responses as opposed to merely airing surveys or rants. **(b)** The criticism distribution shows that help-seekers are more likely to engage in self-criticism (43.32%), and those who criticise openly on others with proper reasoning are more likely to receive assistance. **(c)** The readability statistics of posts in BeCOPE state that well-written posts receive 2.2× more support (responses) as compared to poorly written posts.

https://doi.org/10.1371/journal.pone.0316906.g003

reasoning in criticism is vital for receiving help. In contrast, CNR is most prevalent in the isolated engagement, highlighting that criticism without proper reasoning only adds noisy understanding to the reader's mind. Similarly, *self-criticism* is considered the most prevalent type of criticism among those who receive help. This implies that people seeking support are 20.50% more likely to engage in self-criticism, and those who express their emotions openly are 16.00% more likely to receive assistance. As a result, we infer that peers who criticise and have

a profound comprehension of the topic at hand are more apt to receive assistance. The four annotated criticism labels receive adequate agreement score with confidence $\geq$95% on the *p*-values of *criticism w/ reasoning* (0.043), *criticism w/ no reasoning* (0.010), *no criticism* (0.009), and *self-criticism* (0.035).

3.1.3 Readability. We hypothesize that well-written posts (i.e., easier to read) foster better understanding and subsequently attract more peers to engage. Our initial observation supports the hypothesis; most of the posts in the BeCOPE dataset are hard to read, i.e., rated ≤ 2 on a scale of 1 to 5, with 1 being the least comprehensible. Our analyses reveal that posts scoring higher in readability result in 2.2× greater support ratings from peers, as shown in Fig 3(c). We further employ experts in linguistics to understand what contributes more toward understanding posts. We observe that factors like the length of the post, the division into paragraphs and listicles, grammar, spelling, clarity of the issue, and usage of short forms (SMS language) are critical that peers take into consideration when reading and deciding to engage with a post. The readability score receives significant confidence of $\geq$95% with average *p*-values across all five labels to be 0.040.

3.2 RQ2: How does the expression of emotions in posts impact user engagement in the OMHC platforms?

Emotion labels. Emotions play a vital role in mental health support seeking. Empathetic understanding is an attempt by the observers/experts to regulate emotions that help-seekers express [22]. Fig 4 shows a frequency-based radial distribution of the most frequent emotion labels in BeCOPE. Our analysis of emotion labels shows that 10% of the isolated posts carry *neutral* emotion labels. In contrast, only 3% posts carry *neutral* emotions for both interactive and non-interactive posts combined. Furthermore, 12.3% of the non-isolated posts exhibit *curiosity* as the secondary emotion compared to 7% isolated posts. Evidently, labels such as *sadness, curiosity, fear,* and *realization* are more prevalent in non-isolated posts. On the other hand, emotion labels such as *caring, confusion, approval, joy,* and *neutral* are more prevalent in isolated posts. Consequently, peers exhibiting explicit emotional expression in posts, such as curiosity, fear, and sadness, receive more significant support in 86% of the cases. For the remaining 14% of the posts, emotions are observed to be with tepid emotional labels, such as caring, confusion, or neutral, to which peers often ignored responding, leading to no interaction.

On analyzing a sample of 100 posts, we subjectively categorize extreme emotions expressed into various types, including fear, excitement, sadness, etc. In the category-wise emotion distribution (Fig 4), we observe that posts expressing such explicit extreme emotions have a higher chance of receiving a response, whereas posts with tepid emotional labels, such as caring, confusion, and neutral tend to be ignored.

3.3 Metadata and content analysis

We conduct an auxiliary analysis of the BeCOPE dataset with a prime focus on metadata and textual properties. These experiments aim to assess the impact of minor actions, such as subjectivity, interaction count, anonymity, etc., on help-seeking. We conclude that specific minor actions taken by help-seekers on OMHC platforms can increase the probability of receiving assistance. Our initial findings suggest that descriptive titles and body content attract help, whereas compact usage of words avoids help. Likewise, the active participation of the help-seeker in the conversation (through comments) increases the chances of receiving help by 100%. Such approaches might assist help-seekers in gaining early access to assistance. We observe that a few mental health subreddit channels like *r/OpiatesRecovery* are entirely
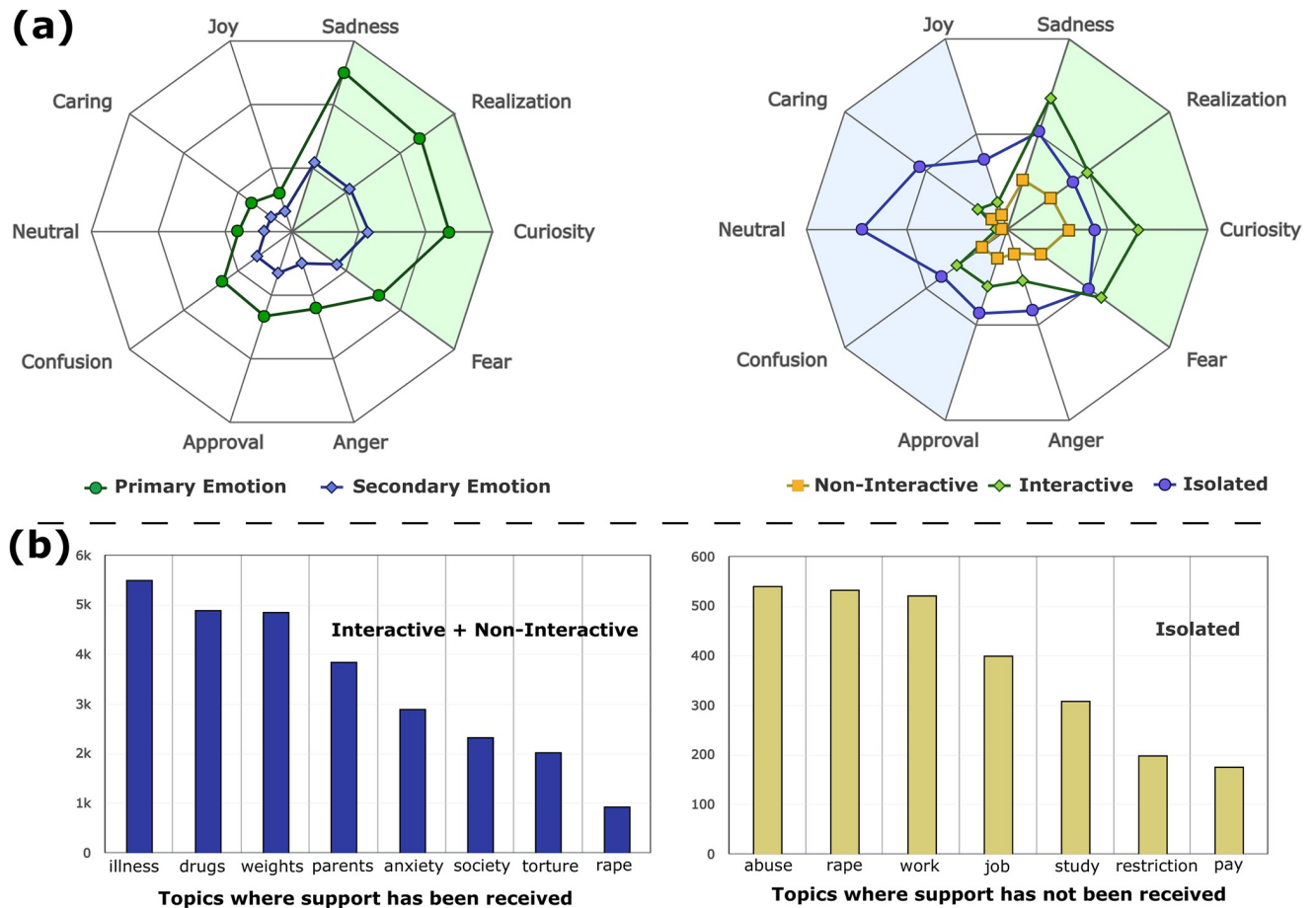
Fig 4. (a) Distribution of emotion labels in the BeCOPE dataset. For brevity, we show plots for the top 10 emotion labels only. Each post is tagged with primary and secondary emotion labels. We further analyze the emotion label distribution across three engagement categories. **(b) Topical analysis on the BeCOPE dataset**. We perform Latent Dirichlet Allocation (LDA) [23] to form 8 clusters of topics. To analyze the topics on which peers respond, we club interactive and non-interactive posts, where peers respond and compare them with topics from isolated posts.

https://doi.org/10.1371/journal.pone.0316906.g004

dedicated to providing frequent assistance to help-seekers, including during late hours. A detailed analysis with additional experiments is presented in S1 Appendix (Section 5).

3.4 Topical analysis

We also perform a topical analysis of peer-to-peer interactions, aiming to understand what specific topics and keywords drive the conversation in three engagement categories (viz. interactive, non-interactive, and isolated). To this end, we apply Latent Dirichlet Allocation (LDA) [23] on the posts in each engagement category. The idea is to understand the topics on which peers respond and don't respond. Therefore, we segregate isolated and non-isolated posts to study the topics on which the support is received and not received, respectively. We observe that the most common topics are in isolated posts, which include discussions about school-related issues, abuse, rape, pressure to meet society's standards, salary, and freedom to express opinions and feelings. On the other hand, we observe that the frequently discussed topics from the non-isolated category are anxiety, drugs, common symptoms/illness and diagnosis, parenting behaviors, body image issues, food and weight, anxiety, and relapsing on drugs. Fig 4

shows a cluster of topics for posts from each category to obtain the most common topics in conversations. Evidently, the common topics of discussion in isolated posts elucidate that people shared experiences about many sensitive and stigmatized issues; subsequently, they remain unexplored, as indicated by the number of isolated posts. As a result, common topics that resonate with peers and enjoy widespread prevalence tend to attract relatively more interactions and active engagement from peers on OMHCs.

4 Discussion

Understanding user behavior and online engagement is consistently challenging, particularly in comprehending the complexities of individuals in distress. OMHC platforms have emerged as crucial spaces for peer-based mental health discussions, enabling individuals to discuss their intrinsic thoughts and mental health issues openly. Beyond the OMHC's function, only a handful of these users interact, with even fewer users receiving the anticipated assistance. The most effective way of assessing peer engagement is to understand the factors on which peer interaction depends. Platforms like Reddit, containing dedicated mental health subreddits, offer rich repositories of discussions on relevant topics. Our formulated hypothesis posits that the comprehension of peer behavioral attributes such as intent, criticism, and readability significantly contributes to a holistic understanding. In addition, the expressivity of emotions on OMHCs can further concentrate on the causal underpinnings of these behavioral dynamics. However, this research area has remained under-resourced and insufficiently explored. Our newly introduced BeCOPE dataset holds significant implications beyond the insights drawn in this study. It can serve as a valuable resource across various research domains with dimensions ranging from empathetic to behavioral conduct of peers on OMHCs and further epitomizing explanations and casualty of such implicit underlying causes.

Our research examines the behavioral, emotional, and topical dynamics associated with varying levels of engagement among peers within OMHCs. We perceive engagement as an indication of a peer's preparedness to provide support. Our findings underscore that simple behavioral characteristics such as explicitly seeking help and refraining from criticizing others can increase peer engagement, as observed in $\sim 50\%$ of the cases. This observation emphasizes that behaviors like ranting, criticising others, and generic chit-chatting do not elicit productive peer attention. At the same time, users express themselves in different styles, and the underlying concept of peers being able to understand others hinges on the clarity of the posts' readability. Earlier research shows that using short sentences is more engaging [24]. In contrast, we show that peers with intricate thoughts aren't constrained to concise posts; instead, they often require more extensive elaboration [25]. Our research demonstrates a twofold increase in support for individuals openly expressing their concerns on the OMHC platforms. Conversely, the illustration of emotion dynamics is an additional gauge to evaluate the user's context. In alignment with our formulated hypothesis, the intricate interplay of emotions articulated within OMHCs demonstrated a direct correlation with the level of peer interaction. Analogous to socio-cultural implications, instances where individuals convey heightened emotional intensity consistently involve more engagement, while expressions characterized by emotional neutrality tend to diminish in terms of peer involvement. This phenomenon potentially stems from underlying factors such as relatability, the emergence of a palpable sense of urgency, and a compelling inclination to provide empathetic validation and support. These emotionally charged interactions establish a conspicuously relatable presence, effectively motivating peers to participate in discussions. Consequently, the assessment of peer engagement within OMHCs stands as pertinent societal research that aims to assess the intricate dynamics underpinning an effective peer support framework. Such OMHCs serve as forums where peers

engage in a wide spectrum of discussions, yet only a few receive the required assistance. We are convinced that a crucial void in this landscape lies in fostering societal awareness regarding the nature of these challenges and their appropriate navigation. For instance, individuals often discuss sensitive and stigmatized matters, which, although prevalent in volume, remain relatively unexplored, as substantiated by the prevalence of isolated posts. As a result, topics of a general nature are observed to attract increased interaction. Furthermore, there exist a few impactful takeaways from our auxiliary content (metadata) analysis. We present a detailed discussion in S1 Appendix. These perceptive insights inherently underscore the significance of understanding the factors of the support ecosystem before its effective utilization for constructive engagement.

5 Conclusion

OMHC platforms have become a popular way to seek help for people struggling with mental health issues [26–29]. Our work analyzed the granular user posting behaviors that foster peer engagement with the mental health content on OMHC platforms, specifically subreddits. The primary aim of this work was to better understand the behaviors of support seekers and the factors that drive peer engagement with the original post. We found that the intent of a post (seeking support versus ranting about one's experience), the readability, and the criticism elements of a post were associated with peer engagement. Our proposed dataset and empirical study call for more research to understand peer engagement on mental health platforms, including elements that lead to constructive versus detrimental engagement [28, 30, 31]. These data are critical in understanding how OMHC can best support users experiencing distress in addition to preventing the proliferation of harmful and inaccurate mental health advice and information [32–34].

Understanding user behavior and online activity is challenging, and even harder to understand individuals in distress. The current study primarily focused on peer-to-peer engagement concerning mental health content. We understand that the findings can vary across other platforms like Twitter, Talklife, 7Cups, Facebook, Instagram, and even other subreddit channels. The future direction of this work will be to better understand user behavior on OMHCs, including how to monitor and moderate peer engagement so that it is not harmful to individuals in distress. Although our findings shed light on the connecting patterns of peer-to-peer online engagement, more research is needed to develop computational methods to gauge user satisfaction and behavior by exploiting the annotations we have done in BeCOPE.

Supporting information

S1 Fig. Mental health subreddits and their frequencies. Distribution of mental health subreddits across all engagement categories.
(PDF)

S1 Appendix. Auxiliary analysis. Attached is an additional section on a set of auxiliary analysis sections discussing the findings on metadata.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Aseem Srivastava, Tanya Gupta.

Data curation: Aseem Srivastava.

Formal analysis: Aseem Srivastava.

Investigation: Aseem Srivastava, Tanya Gupta.

Methodology: Aseem Srivastava.

Project administration: Md Shad Akhtar.

Resources: Alison Cerezo, Sarah Peregrine (Grin) Lord, Md Shad Akhtar.

Supervision: Alison Cerezo, Sarah Peregrine (Grin) Lord, Md Shad Akhtar, Tanmoy Chakraborty.

Validation: Aseem Srivastava.

Writing – original draft: Aseem Srivastava, Tanya Gupta.

Writing – review & editing: Alison Cerezo, Sarah Peregrine (Grin) Lord, Md Shad Akhtar, Tanmoy Chakraborty.

References

1. Butryn T, Bryant L, Marchionni C, Sholevar F. The shortage of psychiatrists and other mental health providers: Causes, current state, and potential solutions. International Journal of Academic Medicine. 2017; 3(1):5–9. https://doi.org/10.4103/IJAM.IJAM_49_17

2. Rathod S, Pinninti N, Irfan M, Gorczynski P, Rathod P, Gega L, et al. Mental Health Service Provision in Low- and Middle-Income Countries. Health Services Insights. 2017; 10:1178632917694350. https://doi.org/10.1177/1178632917694350 PMID: 28469456

3. Lokala U, Srivastava A, Dastidar TG, Chakraborty T, Akhtar MS, Panahiazar M, et al. A Computational Approach to Understand Mental Health from Reddit: Knowledge-Aware Multitask Learning Framework. Proceedings of the International AAAI Conference on Web and Social Media. 2022;16(1):640–650.

4. Park A, Conway M, Chen AT. Examining Thematic Similarity, Difference, and Membership in Three Online Mental Health Communities from Reddit. Comput Hum Behav. 2018; 78(C):98–112. https://doi.org/10.1016/j.chb.2017.09.001 PMID: 29456286

5. De Choudhury M, De S. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. Proceedings of the International AAAI Conference on Web and Social Media. 2014;8(1):71–80.

6. Sharma A, Choudhury M, Althoff T, Sharma A. Engagement Patterns of Peer-to-Peer Interactions on Mental Health Platforms. In: Proceedings of the International AAAI Conference on Web and Social Media. vol. 14; 2020. p. 614–625.

7. Pruksachatkun Y, Pendse SR, Sharma A. 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. CHI'19. New York, NY, USA: Association for Computing Machinery; 2019. p. 1–13. Available from: https://doi.org/10.1145/3290605.3300294.

8. Ernala SK, Labetoulle T, Bane F, Birnbaum ML, Rizvi AF, Kane JM, et al. Characterizing Audience Engagement and Assessing Its Impact on Social Media Disclosures of Mental Illnesses. Proceedings of the International AAAI Conference on Web and Social Media. 2018;12(1).

9. Evans MK, Donelle L, Hume-Loveland L. Social support and online postpartum depression discussion groups: a content analysis. Patient education and counseling. 2012; 87 3:405–10. https://doi.org/10.1016/j.pec.2011.09.011 PMID: 22019021

10. Rains S, Peterson E, Wright K. Communicating Social Support in Computer-mediated Contexts: A Meta-analytic Review of Content Analyses Examining Support Messages Shared Online among Individuals Coping with Illness. Communication Monographs. 2015; 82:1–28. https://doi.org/10.1080/03637751.2015.1019530

11. Andalibi N, Ozturk P, Forte A. Sensitive Self-Disclosures, Responses, and Social Support on Instagram: The Case of Depression. In: Proceedings of the 2017 ACM Conference on Computer Supported

Cooperative Work and Social Computing. CSCW'17. New York, NY, USA: Association for Computing Machinery; 2017. p. 1485–1500. Available from: https://doi.org/10.1145/2998181.2998243.

12. Donnelly AP. A global mental health approach to depression in adolescence. Nat. Mental Health. Nature Mental Health. 2023; p. 527–528. https://doi.org/10.1038/s44220-023-00107-y

13. Seiferth C, Vogel L, Aas B. How to e-mental health: a guideline for researchers and practitioners using digital technology in the context of mental health. Nature Mental Health. 2023; p. 542–554. https://doi.org/10.1038/s44220-023-00085-1

14. Hoven M, Luigjes J, Denys D. How do confidence and self-beliefs relate in psychopathology: a trans-diagnostic approach. Nature Mental Health. 2023; p. 337–345. https://doi.org/10.1038/s44220-023-00062-8

15. Re'em Y, Stelson EA, Davis HE. Factors associated with psychiatric outcomes and coping in Long COVID. Nature Mental Health. 2023; p. 361–372. https://doi.org/10.1038/s44220-023-00064-6

16. Ali K, Farrer L, Gulliver A, Griffiths KM. Online Peer-to-Peer Support for Young People With Mental Health Problems: A Systematic Review. JMIR Mental Health. 2015; 2(2):e19. https://doi.org/10.2196/mental.4418 PMID: 26543923

17. Biagianti B, Quraishi SH, Schlosser DA. Potential Benefits of Incorporating Peer-to-Peer Interactions Into Digital Interventions for Psychotic Disorders: A Systematic Review. Psychiatric Services. 2018; 69 (4):377–388. https://doi.org/10.1176/appi.ps.201700283 PMID: 29241435

18. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiology and Psychiatric Sciences. 2016; 25(2):113–122. https://doi.org/10.1017/S2045796015001067 PMID: 26744309

19. Kazdin AE. Annual Research Review: Expanding mental health services through novel models of intervention delivery. Journal of Child Psychology and Psychiatry. 2019; 60(4):455–472. https://doi.org/10.1111/jcpp.12937 PMID: 29900543

20. Yan H, Fitzsimmons-Craft EE, Goodman M, Krauss M, Das S, Cavazos-Rehg P. Automatic detection of eating disorder-related social media posts that could benefit from a mental health intervention. International Journal of Eating Disorders. 2019; 52(10):1150–1156. https://doi.org/10.1002/eat.23148 PMID: 31381168

21. Reimers N, Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics; 2019. Available from: https://arxiv.org/abs/1908.10084.

22. Kim JS, Franklin C. Understanding emotional change in solution-focused brief therapy: Facilitating positive emotions. Best Practices in Mental Health. 2015; 11(1):25–41.

23. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. Journal of machine Learning research. 2003; 3 (Jan):993–1022.

24. Althoff T, Clark K, Leskovec J. Large-scale Analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health. Transactions of the Association for Computational Linguistics. 2016; 4:463–476. https://doi.org/10.1162/tacl_a_00111 PMID: 28344978

25. Sit M, Elliott SA, Wright KS, Scott SD, Hartling L. Youth Mental Health Help-Seeking Information Needs and Experiences: A Thematic Analysis of Reddit Posts. Youth & Society. 0;0 (0):0044118X221129642.

26. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The Effectiveness of an Online Support Group for Members of the Community with Depression: A Randomised Controlled Trial. PLOS ONE. 2012; 7(12):1–9. https://doi.org/10.1371/journal.pone.0053244 PMID: 23285271

27. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally Occurring Peer Support through Social Media: The Experiences of Individuals with Severe Mental Illness Using YouTube. PLOS ONE. 2014; 9 (10):1–9. https://doi.org/10.1371/journal.pone.0110171 PMID: 25333470

28. Odgers CL, Jensen MR. Annual Research Review: Adolescent mental health in the digital age: facts, fears, and future directions. Journal of Child Psychology and Psychiatry. 2020; 61(3):336–348. https://doi.org/10.1111/jcpp.13190 PMID: 31951670

29. Elmquist D, McLaughlin C. Social Media Use Among Adolescents Coping with Mental Health. Contemporary School Psychology. 2017; 22.

30. Gan DZQ, McGillivray L, Larsen ME, Torok M. Promoting engagement with self-guided digital therapeutics for mental health: Insights from a cross-sectional survey of end-users. Journal of Clinical Psychology. 2023; 79(5):1386–1397. https://doi.org/10.1002/jclp.23486 PMID: 36693234

31. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. BMJ Ment Health. 2018; 21(3):116–119. https://doi.org/10.1136/eb-2018-102891 PMID: 29871870

32. Palmer K, Burrows V. Ethical and Safety Concerns Regarding the Use of Mental Health–Related Apps in Counseling: Considerations for Counselors. Journal of Technology in Behavioral Science. 2021; 6. https://doi.org/10.1007/s41347-020-00160-9 PMID: 32904690

33. Burr C, Morley J, Taddeo M, Floridi L. Digital Psychiatry: Risks and Opportunities for Public Health and Wellbeing. IEEE Transactions on Technology and Society. 2020; 1(1):21–33. https://doi.org/10.1109/TTS.2020.2977059

34. Oyebode O, Alqahtani F, Orji R. Using Machine Learning and Thematic Analysis Methods to Evaluate Mental Health Apps Based on User Reviews. IEEE Access. 2020; 8:111141–111158. https://doi.org/10.1109/ACCESS.2020.3002176