



# The Role of Moderators in Facilitating and Encouraging Peer-to-Peer Support in an Online Mental Health Community: A Qualitative Exploratory Study

Davy Deng<sup>1</sup> · Tim Rogers<sup>2</sup> · John A. Naslund<sup>3</sup>

Received: 23 October 2022 / Revised: 4 December 2022 / Accepted: 17 January 2023 / Published online: 16 February 2023  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

## Abstract

Online peer support platforms have gained popularity as a potential way for people struggling with mental health problems to share information and provide support to each other. While these platforms can offer an open space to discuss emotionally difficult issues, unsafe or unmoderated communities can allow potential harm to users by spreading triggering content, misinformation or hostile interactions. The purpose of this study was to explore the role of moderators in these online communities, and how moderators can facilitate peer-to-peer support, while minimizing harms to users and amplifying potential benefits. Moderators of the Togetherall peer support platform were recruited to participate in qualitative interviews. The moderators, referred to as ‘Wall Guides’, were asked about their day-to-day responsibilities, positive and negative experiences they have witnessed on the platform and the strategies they employ when encountering problems such as lack of engagement or posting of inappropriate content. The data were then analyzed qualitatively using thematic content analysis and consensus codes were deduced and reviewed to reach final results and representative themes. In total, 20 moderators participated in this study, and described their experiences and efforts to follow a consistent and shared protocol for responding to common scenarios in the online community. Many reported the deep connections formed by the online community, the helpful and thoughtful responses that members give each other and the satisfaction of seeing progress in members’ recovery. They also reported occasional aggressive, sensitive or inconsiderate comments and posts on the platform. They respond by removing or revising the hurtful post or reaching out to the affected member to maintain the ‘house rules’. Lastly, many discussed strategies they elicit to promote engagement from members within the community and ensure each member is supported through their use of the platform. This study sheds light on the critical role of moderators of online peer support communities, and their ability to contribute to the potential benefits of digital peer support while minimizing risks to users. The findings reported here accentuate the importance of having well-trained moderators on online peer support platforms and can guide future efforts to effectively train and supervise prospective peer support moderators. Moderators can become an active ‘shaping force’ and bring a cohesive culture of expressed empathy, sensitivity and care. The delivery of a healthy and safe community contrasts starkly with non-moderated online forums, which can become unhealthy and unsafe as a result.

**Keywords** Peer support · Digital mental health · Online mental health community · Moderator · De-stigmatization · Thematic analysis

✉ Davy Deng  
davy@hst.harvard.edu

Tim Rogers  
tim.rogers@togetherall.com

John A. Naslund  
John\_Naslund@hms.harvard.edu

<sup>1</sup> Harvard Chan School of Public Health, Boston, MA, USA

<sup>2</sup> Togetherall, London, UK

<sup>3</sup> Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA

## Introduction

Given the widespread reach of mobile devices and popularity of social media, digital platforms have become avenues through which individuals struggling with mental health problems frequently turn to seek support or find information about their mental health concerns (Peek et al., 2015; Torous et al., 2020). As a result of the coronavirus (COVID-19) pandemic, more people struggling with mental health problems turned to online peer support groups for counseling,

coaching and seeking advice (Stein et al., 2022). Among many forms of digital mental health, peer support programs have been shown to facilitate information exchange for medication, coping skills, peer therapy, storytelling and emotional support (Skousen et al., 2020; Tanis, 2008). Extensive research has shown that for some individuals, use of these asynchronous online peer support platforms is associated with similar benefits of in-person talk therapies, including increased feeling of connectedness and hope, enhanced self-efficacy and self-esteem, as well as an elevated sense of belonging (Bickerstaff et al., 2021; Naslund & Deng, 2021; Prescott et al., 2020). Other studies have reported an increase in supportive communication and emotional well-being, and beneficial psychosocial outcomes for patients with severe psychiatric disorders (Fortuna et al., 2020; Naslund et al., 2018). Furthermore, a recent review found that digital peer support platforms are feasible and hold strong potential for achieving clinical effectiveness (Fortuna et al., 2020). They also have great potential for public health impact as they are also highly scalable, and can reach many more people where traditional in-person approaches to peer support or group interventions cannot, especially in low- and middle-income countries where effective resources for mental health support are scarce (Naeem et al., 2020; Naslund & Deng, 2021).

However, in contrast, there are also studies that have described the challenges with online peer support platforms, and the limited impact that these platforms can achieve (Griffiths et al., 2015), often due to high attrition rates among users of the platforms, especially those without active engagement methods or where mental health professionals are not continuously on standby to provide consultation when needed (Eysenbach, 2005; Finnerty et al., 2019; Salyers et al., 2017). Some studies even observed an increase in mental health stigmatization associated with more cyberbullying and reinforcement from unsafe online communities without effective moderation (Heinsch et al., 2021; Martinez-Martin & Kreitmair, 2018; Yew, 2021). Therefore, despite the wide audience and increasing reach of online mental health communities, it is important to understand how peer support programs can facilitate and ensure a quality experience for their users, bring safe and beneficial impact to users' mental health, while minimizing negative impacts.

In particular, programs with dedicated moderators to manage content, ensure safety, and support users when needed, appear to be successful by achieving lower attrition rates among users (Salyers et al., 2017). The extent of mental health training, professional background and registration/accreditation of moderators of these groups is a key consideration. They can then play an essential role in supporting positive interactions on these platforms and promoting engagement among users (Fortuna et al., 2020). Because of their close proximity to the

target population and their role of ensuring the quality of interactions on these online platforms, as well as promoting safety on the site, moderators could potentially offer valuable insights about their experiences, as well as suggestions for how to improve the quality of online peer support programs. They may also be ideally positioned to comment on how to address the potential challenges of online support groups such as responding to risk issues among users (such as potential suicide, domestic violence or child safety concerns), addressing the spread of harmful or misleading content, responding to hostile interactions among users, as well as promoting user engagement and preventing attrition. Their experience could also shed light on the core components and characteristics that render online peer support programs beneficial for users, while simultaneously yielding potential solutions to the challenges that limit the uptake and impact of these platforms. For instance, the moderators of online peer support groups could provide insights necessary to overcome low user engagement, as well as addressing potential concerns about the lack of truly safe spaces to discuss sensitive matters. The insights from moderators could advance efforts to ensure online peer support can be further expanded and strengthened as one of the mainstays of informal mental health support (Cataldo, 2009; Kraut et al., 2011). However, few studies have considered the perspectives of moderators of online peer support groups, and the ways in which their perspectives could support efforts to better understand how to improve the user experience on these platforms.

Thus, this study explored the perspectives of the moderators of a popular international online peer support platform towards informing efforts to improve user experience and promote engagement in online peer support among users. This study employed a qualitative approach, consisting of in-depth focused interviews to address the following objectives. The first objective primarily explored perspectives of the moderators of a popular online peer support group for mental health including their training, continuous on-the-job support, strategies and techniques they employ to promote user engagement and safety, and overview of their daily activities and responsibilities. The second objective aimed to better understand from the perspective of moderators' what recommendations/strategies they may have for improving the quality of interactions in online peer support groups to promote engagement among users and to create a safe online environment for offering support to persons experiencing mental health problems. As part of this study, participants were encouraged to comment based on their own experiences in their role as moderators. The overarching goal of this study was to inform future online peer support initiatives and to guide the use of better strategies to enhance the experience of the users and of the moderators. This is particularly important because online peer support platforms represent a

potentially scalable approach to addressing population-level mental health challenges, further emphasizing the need to ensure that these platforms are safe and that strategies are implemented to promote the benefits while minimizing risks to users.

## Materials and Methods

### The Togetherall Platform

In this study, we recruited moderators of the Togetherall online peer support platform (<https://Togetherall.com/en-us/>). Togetherall is a digital mental health platform with more than a decade of experience, through which users at more than 500 institutions and organizations globally seek mental health care through directed peer support. The platform was formerly called Big White Wall and was first launched in the UK in 2007, expanded to Australasia in 2011 and the USA in 2014 (Hensel et al., 2016; Morriss et al., 2021). It is a platform open for anyone to register and join at no cost. Users can freely post their mental health struggles and comment on other people's posts to offer emotional support and tips for coping. There are also topic-specific affinity groups formed both by the platform or by users for more thorough discussions. It provides mental health learning, courses and self-report clinical tests that are also moderated for safety.

The Togetherall platform relies on professionally registered/licensed mental health practitioners to moderate, called "Wall Guides". They ensure a positive and validating environment for members who belong to the community. Wall Guides are all trained in counseling, social work or related fields prior to joining the Togetherall platform. The majority of them are professionals with an advanced degree in a mental health-related field (such as counseling psychology and social work). Wall Guides are supervised by experienced 'Lead Wall Guides' who, in turn, receive supervision around the clock from senior clinicians such as an on duty consultant psychiatrist (MD). Subsequent to joining Togetherall, Wall Guides are also given shadowing experience supplemented with regular seminars and participate in regular check-ins with lead Wall Guides to ensure their ability to facilitate effective peer support on the platform. Wall Guides continuously identify at-risk individuals, with the assistance from artificial intelligence (AI) for harm reduction, remove triggering or hostile comments or posts from the community space (such as overly explicit or sexual content, specific descriptions of the methods used for self-harm or anything that is potentially emotionally charged), encourage user engagement to support each other, and maintain the anonymity of every user by renaming or rewording any potentially identifying information posted on the platform. When a member posts in the community about issues of imminent risk (such as plans to harm themselves or others), on duty senior clinicians (such as a consultant psychiatrist/MD)

escalate the incident to clinical and emergency services such that local face to face intervention can be initiated (for example a police welfare check to locate a member). The Wall Guides also ensure that members do not form overly tight bonds with any singular member or with the Wall Guides themselves, which can lead to potentially unhealthy dependency, thus ensuring consistent utilization of peer support between community members.

### Participants

We received a list of all active Wall Guides and lead Wall Guides (44 in total, lead Wall Guides have additional supervising duties for the newer Wall Guides) from the Togetherall platform leadership team and invited them to participate via email in this exploratory study. A total of 20 Wall Guides replied to our initial email outreach to express their interest. The names and information of participants, along with their responses were strictly kept from anyone affiliated with Togetherall to prevent possible influence by the Togetherall policies and procedures. An online consent form was then sent by email and to review the study's aims, methods and data protection procedures. We emphasized to all participants that the study was strictly anonymous and would not impact their standing as employees of Togetherall. We also ensured that only the Harvard-affiliated researchers on our team would conduct the interviews and that there would be no Togetherall team involvement in the entire data collection or data interpretation and analysis process to minimize risk of potential bias. Participants who gave consent were then prompted to complete a short survey with basic demographics questions using a link to Qualtrics online survey software. While we had intended the interviews to be group discussions, due to time zones and differences in availability, we eventually conducted all but one interview in a one-on-one setting over Zoom videoconferencing software. We set the data collection time frame over a 2-month period from August 14th to October 14th, 2021.

### Ethical Issues

Approval was obtained prior to commencement of the project from the Harvard Medical School and Harvard T.H. Chan School of Public Health Institutional Review Boards (IRB) for research with human subjects. We emphasized to all participants before the start of the qualitative interviews that their participation is entirely voluntary, and that they can choose to stop the interview at any time. We also ensured data privacy and that anything brought up during the interview was kept strictly confidential. Only de-identified aggregate data was considered for inclusion in the analysis and interpretation of the findings in preparation for publication. Personally identifiable data were completely removed from

all quotations to ensure anonymity. Participants were given 50 euros each as a token of appreciation for their time following completion of the interview.

## Data Collection

Interviews lasted approximately one hour, and all interviews were conducted by a graduate student from the Harvard Chan School of Public Health and faculty member at Harvard Medical School. The interviews were semi-structured and were accompanied by a fifteen-question semi-structured interview guide. The guide included high-level questions that covered main topics of discussions including the Wall Guide's challenging and rewarding experience moderating discussions on the Togetherall platform, as well as their commonly used strategies or opinions about how to ensure a safe and supportive community. The interviews followed the guide, although were not restricted to the specific questions, where additional probes and open discussion were encouraged to further explore interesting topics. Data was collected through teleconference software (e.g., Zoom) via audio-recording and supplemented by the interviewer's type-written field notes. The audio recordings were transcribed using the auto transcription feature in the Zoom video-conferencing software. Both recordings and audio transcripts were kept in a secure encrypted folder housed on the Harvard University servers. Demographic data were collected through anonymous surveys sent through Qualtrics and aggregated results were also kept in the secure Harvard server.

## Data Analysis

We employed a deductive approach to thematic content analysis (Bradley et al., 2007), which started with the analysis from a broad, organizational framework of coding category based on the interview questions. Two coders were a graduate student in the Harvard T.H. Chan School of Public Health (DD) and an instructor at Harvard Medical School (JN), respectively. To begin the thematic analysis, both coders familiarized themselves with the entire data set by reading and rereading audio-generated transcripts. The lead interviewer (DD) generated a list of initial codes by referring to the field notes and observations from interviews and deducing from the broad topics covered in the original interview guide. We also referred to existing social support theories when constructing codes for analysis. We considered these theories to help frame the survey questions and inform our interpretation of the meaning of the findings, as these theories can help to explain the relationships between engaging in an online peer support community, deriving support from these online interactions, and the resulting impact on health and wellbeing (Wright, 2016). For instance, we considered the social information processing theory, which

suggests hyperpersonal interactions can form when the relationship is constricted to a virtual environment and can often be beneficial for high quality peer support, in considering the impact of the virtual/online format on peer relationships (Wright, 2000). Additionally, we applied the strength of weak tie theory, which postulates that individuals tend to seek social support through weak ties rather than strong ties due to diverse opinions and information, and to maintain anonymity, in order to inform our understanding of progress made by members supported by anonymous strangers on the platform (Wright & Miller, 2010).

After generating the initial code list, a random selection of 2 transcripts was then separately analyzed line-by-line by each coder. Additional codes were assigned by the coders when a new concept became apparent through line-by-line review of the transcript. This allowed more codes to be supplemented to the initial list. The two coders then met to review the code list, and to reach consensus before proceeding with coding the remaining transcripts. The lead coder then iteratively applied the process to 2 transcripts at a time and met regularly with the secondary coder to obtain consensus on a revised list through comparison and discussion. As more data were reviewed, the code list was specified and refined to fit the data better. This produced 12 codes in total. The code and code structures are considered complete when it reaches saturation where no new conceptual categories are generated from reviewing additional transcripts (Bradley et al., 2007). The complete code list was then grouped into 3 overarching categories based on content similarities. Lastly, the senior researcher (JN) and the Togetherall clinical director (TR) provided feedback, which resulted in some minor changes to the coding labels.

## Results

### Study Participants

During the 2-month study period, we emailed the survey to all 44 Wall Guides that were at the time active employees of the Togetherall platform. In total, 20 wall guides (45%) completed the anonymous online questionnaire. The mean age of respondents was 40.65 (SD = 8.43) years. The vast majority of respondents were female (90%) and White (90%), and over half reported having extensive experience in psychotherapy with master's degree level training or above. Of the 20 participants, 18 eventually completed interviews, while 2 did not respond after 3 email reminders. It should be noted that our demographics are relatively homogeneous due to the limitation of available sample sets, and cautions should be taken when generalizing the results to other demographics or cultural settings. Detailed demographic and work experience information is summarized in the Table 1 below.

**Table 1** Demographic characteristics of the participating moderators of the Togetherall online peer support platform

	N	%
<b>Years of Experience</b>		
<1 year	3	15
>1 year, <2 years	7	35
>2 years, <5 years	8	40
>5 years	2	10
<b>Age</b> (Mean = 40.65 years, SD = 8.43 years)		
<b>Gender</b>		
Male	1	5
Female	18	90
Prefer not to say	1	5
<b>Race or ethnicity</b>		
Asian	1	5
Caucasian/Non-Hispanic White	18	90
Black/African American	1	5
<b>Highest degree received</b>		
High school graduate or equivalent	1	5
Some college or certificate program	3	15
Bachelor's degree or equivalent	5	25
Master's degree	10	50
PhD or equivalent	1	5
<b>Field in which you received your degree</b>		
Psychology/Psychotherapy/Counseling	7	64
Social work/Social sciences	2	18
Others	2	18
<b>Have you received training related to mental health</b>		
Yes	20	100
<b>Have you had experience delivering any structured psychological therapies (e.g. BA, CBT, IPT or others)</b>		
Yes	13	65
No	7	35
<b>How many years of experience have you had delivering such therapy</b>		
<5 years	6	46
>5 years	7	54
<b>Have you had experience specifically with CBT</b>		
Yes	12	60
No	8	40
<b>Have you had experience specifically with BA</b>		
Yes	17	85
No	3	15
<b>On average how many patients with mental health problems do you interact per month in the past 12 months outside of Togetherall</b>		
<5	7	37
>5, <20	7	37
>20, <40	2	10
>40	3	16
<b>On average how many members of Togetherall do you interact per month in the past 12 months</b>		
Difficult to estimate/No ideas	2	10
<50	5	26
<50, <100	4	21
>100	8	42
<b>Country</b>		

**Table 1** (continued)

	N	%
UK	15	75
Canada	1	5
New Zealand	4	20

## Qualitative Findings

We identified major themes and categories reflecting participants': (1) interpretation of their role as a Wall Guide on the Togetherall platform; (2) top positive experiences moderating online peer-to-peer support; and (3) ways to respond to challenging situations and/or inappropriate behaviors on the platform. Table 2 provides a summary of these major topics, the assigned codes from the coding list, and selected representative quotes from participants. The broad categories are also summarized below:

1. **Key responsibilities of Wall Guides.** Many participants mentioned the distinction between acting as a Wall Guide compared to a therapist. While most of the Wall Guides were licensed therapists, many emphasized that their role on the platform was to chaperone a safe and positive peer support environment and facilitate or encourage more meaningful engagement from members to support them in helping each other. For example, many would post direct questions on posts that had not attracted much attention in order to elicit answers from the community, and occasionally they would reword the post in such a way that a response would become more likely (e.g., make it more succinct or more organized). Many also mentioned the role of safety net, in which the Wall Guides use their counseling expertise to respond quickly to potential hazardous scenarios that could cause imminent danger to any member or that could be triggering to others in the community.
2. **Positive interaction with members.** A number of participants expressed the sense of accomplishment when seeing members making satisfactory progress towards recovery and mentioned feeling inspired by the tremendous amount of resilience some members have demonstrated. Another major positive aspect the Wall Guides brought up often was the positive human connections. Many Wall Guides described the authenticity and genuine nature of peer support from members on the Togetherall platform who have never met each other, and how caring, encouraging, intimate and respectful relationships can be struck up between anonymous members without the need for in-person communications. Lastly, numerous Wall Guides mentioned an important aspect of their work, where they described their active shaping of a platform that offers a destigmatized and judgment-free space for members to discuss difficult or controversial issues via peer support, including topics such as homophobia, psychosis, self-harm and culture-specific taboos. This active 'shaping force' sits in stark contrast to potentially unhealthy or unsafe non-moderated online forums.
3. **Responding to challenging interaction with members.** With respect to the challenges of the platform, the Wall Guides described the strategies they use to manage emotionally triggering issues in the community, such as the (graphic) description of suicidal methods or overly vivid depiction of the symptoms of an eating disorder. The Wall Guides sometimes mentioned the discomfort they can experience if altering members' posts, while acknowledging that it is necessary to do so in instances where these messages might potentially be distressing for other members. The Wall Guides expressed facing challenges when responding to members who exhibit imminent risk of suicidal behaviors or self-harm, members who suffer from eating disorders, and members who disclose very disturbing or trauma-related thoughts. The Wall Guides described the way in which managing these scenarios, while not frequent among members, still requires meticulous attention in order not to distress the rest of the community. At the same time, having to alter certain member's posts to prevent harm to the community is reported to be among the biggest challenges for a Wall Guide. Most find it hard to gauge the right level of modification: seeking to balance staying true to the original post and permitting freedom of expression "allowing the community to breathe"; yet, still avoiding community harm. Another challenging experience that the Wall Guides mentioned included members' demonstration of hostility, such as using harsh or trolling language, or when members' willfully break house rules (for example intentionally posting explicit messages on the platform or asking other member's socially inappropriate questions). When moderating content that breaks house rules, the Wall Guides reported that they would usually flag the post followed by either paraphrasing or editing out certain content of the post, and then send a message to the original poster to let them know that they had to alter the post. When a post does not receive sufficient engagement or attention from community members, Wall Guides will often comment on the post them-

**Table 2** Summary of the major categories, sample codes, and representative quotes from participating Wall Guides

Major categories	Code for each category	Quotes
Key responsibilities of WG	Safety net to members	We moderate the Together All platform and ensure that the interactions between the members are aligned with clinical safety and house policies. (004) We make sure everyone is safe on the wall and are alerted to certain sensitive words in community post, at which case we send direct messages to some members to ensure their safety. We also provide local resources such as crisis service to address any emergencies. (006)
	Guardian of peer support environment	We monitor what's going on in the community, make contributions to posts and intervene when it's needed (self-harm etc.). In general, we make sure everyone gets as much out of the platform as they can. (016) WG is about staying tune what is happening in the community, hovering above the community for signs of distress and encouraging people to participate in the community. (017)
	Facilitating and encouraging engagement	I encourage other members to interact with a post by posing a question like "does anyone else have any advice or suggestions etc." (001) If I get a direct message from a member, I generally respond to the situation but gently guide them to use the community. Reinforce that the platform is a community platform, and they should talk to other members and use the resources on the platform. Make sure they know that this is not a counseling service, and we do not want to encourage that dependency unless they are in immediate risk. (003)
	Challenging (mental health) stigma	Someone wrote about posts about homophobia, they have difficulties accepting people who are not straight, they want to feel differently but couldn't get past it. The members discuss the idea in a very respectful and honest way, and it was a great conversation. (004) A member is concerned with using medications for depression and other members came in and comforted him that taking the depression medications is no different from medications for diabetes (006)
Positive Interaction with members	Satisfactory progress seen in members or a sense of accomplishment from helping at-risk members	For example, someone who was suicidal came back and told us that they wouldn't be here if it weren't for Together All, and now they have a new job now. (001) I can see a tremendous amount of resilience. Some members have gone through lots of traumas but still push forward and take care of themselves and others. (008)
	Learning about human connections from members	I love watching the interactions and friendship members strike up and taught me the importance of relationships in openly supporting each other. (009) I have learned a lot about the depths of human connections. It never ceases to amaze me how much human interaction can exist even when people cannot meet and look into their eyes and yet our members share such intimate things. It makes you feel like its so lovely to be humans. It constantly amazes me. (013) I've learned so much about what it means to listen genuinely. The power of saying "you are not alone" is often forgotten. We often forget just how much human needs to belong and needs to be listened. (017)

**Table 2** (continued)

Major categories	Code for each category	Quotes
Challenging interaction with members & dealing with lack of engagement and inappropriate content	Difficult or emotionally triggering issues	It is challenging when someone is going through suicidal attempts, and not willing to cooperate and provide personal information for us to get them help in that local area (010)  Sometimes issues could be very challenging. For example, one member was sharing things such as feelings resembling pedophilia (and it's something they know it's wrong). We try to allow them to talk about it in a safe environment, but it was a difficult topic to keep the community safe and not trigger anyone. (017)
	Uncomfortable with altering member's posts	On the personal level, it feels quite difficult and uncomfortable to have to change someone's stories particularly if it is trauma related. Sometimes it is a fine line between what is appropriate and what is inappropriate and sometimes members will disagree on the edits of the post. (004)  It can be tricky when you have community posts related to historical or recent abuse/trauma because it's difficult to get mindful of how to keep the voice honest. (008)
	Hostile attitude from members towards the WGs	I struggle with it when the member gets angry. The anonymity doesn't help because people don't feel like they are talking to a real person, and they can say things that are extreme and unkind. (006)  Sometimes it is very difficult to explain privacy rules to people and they can be quite confrontational and rude. (014)
	Member's breaking or misunderstanding of house rules	We want members to engage but also occasionally members may become too close and want to disclose personal-identifying information. This is tricky to keep up with the post because sometimes the members come up with very creative ways to try to communicate identifying information so they can continue this interaction off-line. This is to be discouraged: if we think about in-person supports such as AA, one of the rules there is that members are part of that particular group should not have relationships outside of the community. The community is a safe space, nobody can know how things are outside of the community. (005)  Lots of people sign up with name-related information, and we must change it and sometimes they are very frustrated or angry with the change because they feel like we are violating privacy and autonomy. (012)  Different people have different style of 'peer support' – For example, some believe support should be direct and others think should be gentle. So, when a person is being very honest and direct with their response, many others might think it's unsupportive. And I had to sometimes give them warnings and tell them where we don't find supportive. (013)
	Flag, edit and paraphrase	I usually try to preempt topics that can potentially become problematic (such as sensitive topics related to religion or minority groups), and will flag the posts so we can monitor that post to assess the language quickly and carefully (003)  We respond to the sensitive post by editing it and notify and remind them to follow the house rules. I sometimes also muted and deactivated a member if they are consistently breaching house rules (009)  I usually hide the sensitive post first so that I have some time to decide what to edit. I sometimes will discuss with other Wall Guides on what to change so we have more than one person's opinions before editing especially when the topics are tricky, and the edits might not honor the post's original story. I am personally on the more cautious side and bring in others to discuss often before changing the posts. (011)
Elicit community response	Sometimes we might ask open questions in posts that have not gathered a lot of responses to elicit community engagement. Sometimes if a member is too dependent on private messaging to us Wall Guides, we will empathize and provide a reflection of what they are going through but we have to push them to keep trying to use the community and remind them that Together All is a peer support platform. (009)  One of the things with hesitant members, I try to be empathetic about their nervousness about posting. Then I encourage them to take a look at other people's post or encourage them to post a reply to other folks' posts first. (012)	



selves, ask questions that elicit other responses from the community or jump start a conversation as an attempt to facilitate the peer support process.

## Discussion

During the pandemic, online peer support platforms emerged as an important way for people struggling with mental health problems to exchange information, confide about mutual experiences, provide and receive social support and share their personal struggles and successes. Numerous previous studies have highlighted the potential mental health benefits of participating in online peer support communities, with some studies suggesting that the more frequently a person engages with their online community, the more effective the platform is for their recovery (Merchant et al., 2022). If unsafe and ineffectively moderated, multiple potentially negative effects or limitations of social media-based peer support have been noted, including the dangers of people who may then face hostile or triggering comments, troll accounts, resulting in lack of consistent and effective engagement (Easton et al., 2017). Limited research has been devoted to study how the role of platforms moderated by mental health professionals can minimize or remove these negative impacts experienced among users with mental health problems, while accentuating the therapeutic benefits (Huh et al., 2013).

By interviewing a cohort of the Togetherall platform moderators, we found that most utilized similar strategies to promote engagement and manage harmful content in the online peer support community. While each moderator handled different situations uniquely, they followed a set of established guidelines in response to hostile, threatening or any threads, messages or posts that are considered concerning or sensitive. They do so by flagging, hiding or paraphrasing the posts to reduce inappropriately emotional or triggering content. They remind member about house rules for posting content, and in extreme cases, revoke access privileges for certain repeatedly offensive members. Importantly, the decision to edit or hide a post is almost always taken in consultation with other moderators, as a team, to balance the fine line between defending free expressions and protecting community safety. Moderators also meet frequently, as a team, to reflect together, learn and align their practice. By approaching community moderation in this way, Togetherall Wall Guides create a highly moderated safe platform in which vulnerable help-seekers can receive support in a healthy online environment where the potential for deleterious interactions is minimized.

In addition to their two main roles, moderators also adhere to additional principles to ensure the continuity of community on the platform. First, similar to Alcoholics

Anonymous, the platform follows a strict set of rules for anonymity and only accesses member's personal information in the event that the member might be an imminent danger to themselves or others. Moderators remove any personally identifying information, whether direct (such as name, address and phone number) or indirect (such as city of residence and street names) in either usernames or posts. This is to discourage people from both accidentally or intentionally sharing information that may lead to communications off the platform. Moderators are also conscious that sometimes a member can become overly reliant on another member and communicate mostly or exclusively to each other without using the broader community, and even secretly try to find out each other's identity in order to meet offline. These are potentially dangerous situations for their own safety. If permitted, this would defeat the purpose of having an anonymous online community and are therefore strictly prohibited by moderators.

An overarching intention in providing experienced and professionally trained mental health practitioners as moderators is that all of the processes above result, cumulatively, in the active shaping of a healthy space. This active 'shaping force' brings a cohesive culture of expressed empathy, sensitivity and care that helps to deliver a healthy and safe community. The successful achievement of this kind of anonymous, non-judgemental and supportive space sits in stark contrast to non-moderated online forums, which can become unhealthy and effectively unsafe as a result.

To date, most studies of online peer support platforms for mental health have focused on understanding the experiences of users (Belleville et al., 2019; Bunnell et al., 2017; Moor et al., 2019; Ruggiero et al., 2015; Wagner et al., 2012). Similar to many studies and users reporting advantages or positive perceptions of online support groups, Wall Guides have also expressed their amazement over the tight bond that appears to form between people who are complete strangers and who essentially communicate in an anonymous manner. This was further reflected by the deep and lasting connections and continuous support that members can demonstrate towards each other on the platform. This account of mobilizing perceived and received support from community members would appear clearly to have the potential to effectively buffer stress. However, while the majority of the evidence points to the positive impact of such a platform, other literature reports mixed outcomes. There appears to be a significant portion of users who report lack of effective personal changes if unhelpful social interactions or contact with community members is permitted (Griffiths et al., 2015).

While there are few studies that have considered the role of moderators, several important findings align with our study on the importance of moderator roles. For example, one study looked into an online patient community and showed that common challenges that the platform and

moderators face include promoting member participation, divulging of personal information, offering irrelevant or even dangerous advice, and engaging in heated conversations (Skousen et al., 2020). This is parallel to what the Togetherall Wall Guides reported in the current study, and exactly the problems that moderators are poised to address. Another study expressed challenges in engaging users and showed that moderators tend to lend their emotional support and advice as well on the platform, consistent with what we observed among the interviews with the Wall Guides in Togetherall (Windler et al., 2019). Interestingly, one study also suggested that in addition to the myriad roles moderators play in the support group, they also use forums for their own supportive needs such as sharing their own stories and asking questions that are indistinguishable from those from other users (Smedley & Coulson, 2017). This is not however the practice of moderators in the Togetherall community, which is instead “to allow the community to breathe”, i.e. for members themselves to drive the themes and topics and content being posted about.

While the rapid growth of digital peer support programs has led to an abundance of new opportunities for people living with mental health problems to access support, many of these platforms are unmoderated (such as most social media platforms), and the evidence on the impact and potential benefits of these platforms remains mixed. Without effective moderation, use of online peer support, while promising, could have the unintended consequence where the already-vulnerable population of individuals experiencing mental health problems could be exposed to a large influx of harmful content, which must be balanced against the supportive content (Kaplan et al., 2011; Schrank et al., 2010). Our study is one of the few current studies that looked at the role of professional moderators in chaperoning the online community and examining the specific strategies they utilize to maximize positive outcomes of peer support. While our results should not be taken as direct evidence on the harm of unmoderated platforms, it brings awareness to the importance of having trained, professionally registered mental health practitioners to moderate and safeguard the interactions in online communities. The findings here can provide guidance in developing training programs for mental health peer supporters for future digital mental health programs (Charles et al., 2021).

With the increasing interest in leveraging online peer support, development of novel tools is needed to assist moderators in monitoring and guarding these digital safe harbors (Milne et al., 2019). With the advancement of data-driven precision health and artificial intelligence, there have been attempts to use automated triage to improve and prioritize moderator responsiveness and better protect those most vulnerable (Bickman, 2020; Fiske et al., 2019; Gooding & Kariotis, 2021). Additionally, machine learning algorithms could aid in the decision-making process of moderators and change how

and when an escalation to emergency management is needed to protect the well-being of the users (Graham et al., 2019). Lastly, the legal accountability and ethical implications of online peer support groups, particularly those issues involving online anonymous advice and user privacy, still need to be updated to reflect the rapid expansion of digital peer support groups (Gooding & Kariotis, 2021).

## Limitations

Several limitations should be noted for the interpretation and generalization of the results in this study. First, the sample size of the moderators is relatively small and homogeneous and is restricted to moderators of one online peer support platform, which currently operates primarily in the USA, UK, Canada and New Zealand. It is possible that platforms within a different country would not have worked as effectively due to disparate cultural context. Alternatively, moderators of platforms in other countries or other languages may employ different approaches for promoting member engagement and responding to challenging scenarios. Second, the moderators, while all licensed professionals, have varying years of expertise in working with the digital community and therefore can have different experiences and could employ differing approaches to handle challenges that arise on the platform. Our study was exploratory, and not intended to compare differences between Wall Guides based on their education level or years of experience. Additionally, we also did not interview any members on the platform and how they perceive the role and utility of moderators. Finally, due to the nature of focused interviews and the process of sample recruitment, it is possible that sampling bias occurred and that the moderators that responded and were willing to participate in the study were more likely to agree with organizational guidelines in dealing with risky scenarios on the platform.

## Conclusion

Our study shows that moderators play a potentially critical role and highlights the ways in which they work as one of the shaping forces to maximize the safety and chances of beneficial impact of an online peer support community. The apparent usefulness of effective, continuous moderation could also point future research into examining the potential risks of unmoderated peer communities. While some of these popular platforms might offer promises for expanding access to necessary mental health support and could be considered an adjunct to formal mental health care or part of public health efforts that utilize community resources for improving mental health, there is continued need for research aimed at determining how best to scale up the role of effective moderators to support users and to

realize the benefits of these platforms. Additionally, future studies could include the perspectives of the community members and platform users in the usefulness of moderators, they should also expand on our exploratory study to include other moderated peer support platforms in other sociocultural backgrounds, and languages, particularly in low-and middle-income countries, and finally they could include health areas with more targeted digital self-help groups such as addiction, PTSD, severe mental illness, and eating disorders. The excitement surrounding digital mental health and specifically online peer support in recent years, combined with accelerated demand occurring during the pandemic, will require greater scrutiny and investigation into the key features of these platforms, such as the role of moderators. This will be essential for expanding our understanding of how to both optimize the benefits of these platforms while scaling up access to reach and engage more individuals struggling with mental health problems.

**Author Contribution** All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Davy Deng and John Naslund. The first draft of the manuscript was written by Davy Deng and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

## Declarations

**Ethics Approval and Consent to Participate** This study was performed in line with the Institutional Review Board and approval was granted at Harvard University. Informed consent was obtained from all individual participants included in the study.

**Conflict of Interest** The authors declare no competing interests.

## References

- Belleville, G., Lebel, J., Ouellet, M. C., Békés, V., Morin, C. M., Bergeron, N., Campbell, T., Ghosh, S., Bouchard, S., Guay, S., & Macmaster, F. P. (2019). Resilient - An online multidimensional treatment to promote resilience and better sleep: A randomized controlled trial. *Sleep Medicine*, 64, S214–S215. <https://doi.org/10.1016/j.sleep.2019.11.598>
- Bickerstaff, J. M., Karim, S., Whitman, R. K., Cypher, A., Wiener, S., & Radovic, A. (2021). “You have people here to help you, people like me”: A qualitative analysis of a blogging intervention for adolescents and young adults with depression or anxiety. *Journal of Technology in Behavioral Science*, 6(4), 578–588. <https://doi.org/10.1007/s41347-021-00210-w>
- Bickman, L. (2020). Improving mental health services: A 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, 47(5), 795–843. <https://doi.org/10.1007/s10488-020-01065-8>
- Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: Developing taxonomy, themes, and theory. *Health Services Research*, 42(4), 1758–1772. <https://doi.org/10.1111/j.1475-6773.2006.00684.x>
- Bunnell, B. E., Davidson, T. M., Dewey, D., Price, M., & Ruggiero, K. J. (2017). Rural and urban/suburban families’ use of a web-based mental health intervention. *Telemedicine and E-Health*, 23(5), 390–396. <https://doi.org/10.1089/tmj.2016.0153>
- Cataldo, C. G. (2009). *Cultivating communities of practice: A guide to managing knowledge*, by Eti-enne Wenger, Richard McDermott, & William M. Snyder. Boston, MA: Harvard Business School Press, 2002. 284 pages, hard cover. *Academy of Management Learning & Education*, 8(2), 301–303. <https://doi.org/10.5465/amle.2009.41788855>
- Charles, A., Nixdorf, R., Ibrahim, N., Meir, L. G., Mpango, R. S., Ngakongwa, F., Nudds, H., Pathare, S., Ryan, G., Repper, J., Wharrad, H., Wolf, P., Slade, M., & Mahlke, C. (2021). Initial training for mental health peer support workers: Systematized review and international Delphi consultation. *JMIR Mental Health*, 8(5), e25528. <https://doi.org/10.2196/25528>
- Easton, K., Diggle, J., Ruethi-Davis, M., Holmes, M., Byron-Parker, D., Nuttall, J., & Blackmore, C. (2017). Qualitative exploration of the potential for adverse events when using an online peer support network for mental health: Cross-sectional survey. *JMIR Mental Health*, 4(4), e49. <https://doi.org/10.2196/mental.8168>
- Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research*, 7(1), e11. <https://doi.org/10.2196/jmir.7.1.e11>
- Finnerty, M., Austin, E., Chen, Q., Layman, D., Kealey, E., Ng-Mak, D., Rajagopalan, K., & Hoagwood, K. (2019). Implementation and use of a client-facing web-based shared decision-making system (MyCHOIS-CommonGround) in two specialty mental health clinics. *Community Mental Health Journal*, 55(4), 641–650. <https://doi.org/10.1007/s10597-018-0341-x>
- Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research*, 21(5), e13216. <https://doi.org/10.2196/13216>
- Fortuna, K. L., Naslund, J. A., LaCroix, J. M., Bianco, C. L., Brooks, J. M., Zisman-Ilani, Y., Muralidharan, A., & Deegan, P. (2020). Digital peer support mental health interventions for people with a lived experience of a serious mental illness: Systematic review. *JMIR Mental Health*, 7(4), e16460. <https://doi.org/10.2196/16460>
- Gooding, P., & Kariotis, T. (2021). Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review. *JMIR Mental Health*, 8(6), e24668. <https://doi.org/10.2196/24668>
- Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H.-C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports*, 21(11), 116. <https://doi.org/10.1007/s11920-019-1094-0>
- Griffiths, K. M., Reynolds, J., & Vassallo, S. (2015). An online, moderated peer-to-peer support bulletin board for depression: User-perceived advantages and disadvantages. *JMIR Mental Health*, 2(2), e14. <https://doi.org/10.2196/mental.4266>
- Heinsch, M., Geddes, J., Sampson, D., Brosnan, C., Hunt, S., Wells, H., & Kay-Lambkin, F. (2021). Disclosure of suicidal thoughts during an e-mental health intervention: Relational ethics meets actor-network theory. *Ethics & Behavior*, 31(3), 151–170. <https://doi.org/10.1080/10508422.2019.1691003>
- Hensel, J. M., Shaw, J., Jeffs, L., Ivers, N. M., Desveaux, L., Cohen, A., Agarwal, P., Wodchis, W. P., Tepper, J., Larsen, D., McGahan, A., Cram, P., Mukerji, G., Mamdani, M., Yang, R., Wong, I., Onabajo, N., Jamieson, T., & Bhatia, R. S. (2016). A pragmatic randomized control trial and realist evaluation on the implementation and effectiveness of an internet application to support

- self-management among individuals seeking specialized mental health care: A study protocol. *BMC Psychiatry*, 16(1), 350. <https://doi.org/10.1186/s12888-016-1057-5>
- Huh, J., McDonald, D. W., Hartzler, A., & Pratt, W. (2013). Patient moderator interaction in online health communities. *AMIA ... Annual Symposium Proceedings. AMIA Symposium, 2013*, 627–636.
- Kaplan, K., Salzer, M. S., Solomon, P., Brusilovskiy, E., & Cousounis, P. (2011). Internet peer support for individuals with psychiatric disabilities: A randomized controlled trial. *Social Science & Medicine*, 72(1), 54–62. <https://doi.org/10.1016/j.socscimed.2010.09.037>
- Kraut, R. E., Resnick, P., & Kiesler, S. (2011). *Building successful online communities: Evidence-based social design*. MIT Press.
- Martinez-Martin, N., & Kreitmair, K. (2018). Ethical issues for direct-to-consumer digital psychotherapy apps: Addressing accountability, data protection, and consent. *JMIR Mental Health*, 5(2), e32. <https://doi.org/10.2196/mental.9423>
- Merchant, R., Goldin, A., Manjanatha, D., Harter, C., Chandler, J., Lipp, A., Nguyen, T., & Naslund, J. A. (2022). Opportunities to expand access to mental health services: A case for the role of online peer support communities. *Psychiatric Quarterly*, 93(2), 613–625. <https://doi.org/10.1007/s11126-022-09974-7>
- Milne, D. N., McCabe, K. L., & Calvo, R. A. (2019). Improving moderator responsiveness in online peer support through automated triage. *Journal of Medical Internet Research*, 21(4), e11410. <https://doi.org/10.2196/11410>
- Moor, S., Williman, J., Drummond, S., Fulton, C., Mayes, W., Ward, N., Dovenberg, E., Whitaker, C., & Stasiak, K. (2019). ‘E’ therapy in the community: Examination of the uptake and effectiveness of BRAVE (a self-help computer programme for anxiety in children and adolescents) in primary care. *Internet Interventions*, 18, 100249. <https://doi.org/10.1016/j.invent.2019.100249>
- Morriss, R., Kaylor-Hughes, C., Rawsthorne, M., Coulson, N., Simpson, S., Guo, B., James, M., Lathe, J., Moran, P., Tata, L. J., & Williams, L. (2021). A direct-to-public peer support program (big white wall) versus web-based information to aid the self-management of depression and anxiety: Results and challenges of an automated randomized controlled trial. *Journal of Medical Internet Research*, 23(4), e23487. <https://doi.org/10.2196/23487>
- Naeem, F., Husain, Mo., Husain, Mi., & Javed, A. (2020). Digital psychiatry in low- and middle-income countries post-COVID-19: Opportunities, challenges, and solutions. *Indian Journal of Psychiatry*, 62(9), 380. [https://doi.org/10.4103/psychiatry.IndianJPsychiatry.843\\_20](https://doi.org/10.4103/psychiatry.IndianJPsychiatry.843_20)
- Naslund, J. A., Aschbrenner, K. A., Marsch, L. A., McHugo, G. J., & Bartels, S. J. (2018). Facebook for supporting a lifestyle intervention for people with major depressive disorder, bipolar disorder, and schizophrenia: An exploratory study. *Psychiatric Quarterly*, 89(1), 81–94. <https://doi.org/10.1007/s11126-017-9512-0>
- Naslund, J. A., & Deng, D. (2021). Addressing mental health stigma in low-income and middle-income countries: A new frontier for digital mental health. *Ethics, Medicine and Public Health*, 19, 100719. <https://doi.org/10.1016/j.jemep.2021.100719>
- Peek, H. S., Richards, M., Muir, O., Chan, S. R., Caton, M., & MacMillan, C. (2015). Blogging and social media for mental health education and advocacy: A review for psychiatrists. *Current Psychiatry Reports*, 17(11), 88. <https://doi.org/10.1007/s11920-015-0629-2>
- Prescott, J., Rathbone, A. L., & Brown, G. (2020). Online peer to peer support: Qualitative analysis of UK and US open mental health Facebook groups. *DIGITAL HEALTH*, 6, 205520762097920. <https://doi.org/10.1177/2055207620979209>
- Ruggiero, K. J., Price, M., Adams, Z., Stauffacher, K., McCauley, J., Danielson, C. K., Knapp, R., Hanson, R. F., Davidson, T. M., Amstadter, A. B., Carpenter, M. J., Saunders, B. E., Kilpatrick, D. G., & Resnick, H. S. (2015). Web intervention for adolescents affected by disaster: Population-based randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(9), 709–717. <https://doi.org/10.1016/j.jaac.2015.07.001>
- Salyers, M. P., Fukui, S., Bonfils, K. A., Firmin, R. L., Luther, L., Goscha, R., Rapp, C. A., & Holter, M. C. (2017). Consumer outcomes after implementing commonground as an approach to shared decision making. *Psychiatric Services*, 68(3), 299–302. <https://doi.org/10.1176/appi.ps.201500468>
- Schrank, B., Sibitz, I., Unger, A., & Amering, M. (2010). How patients with schizophrenia use the Internet: Qualitative study. *Journal of Medical Internet Research*, 12(5), e70. <https://doi.org/10.2196/jmir.1550>
- Skousen, T., Safadi, H., Young, C., Karahanna, E., Safadi, S., & Chebib, F. (2020). Successful moderation in online patient communities: Inductive case study. *Journal of Medical Internet Research*, 22(3), e15983. <https://doi.org/10.2196/15983>
- Smedley, R. M., & Coulson, N. S. (2017). A thematic analysis of messages posted by moderators within health-related asynchronous online support forums. *Patient Education and Counseling*, 100(9), 1688–1693. <https://doi.org/10.1016/j.pec.2017.04.008>
- Stein, D. J., Naslund, J. A., & Bantjes, J. (2022). COVID-19 and the global acceleration of digital psychiatry. *The Lancet Psychiatry*, 9(1), 8–9. [https://doi.org/10.1016/S2215-0366\(21\)00474-0](https://doi.org/10.1016/S2215-0366(21)00474-0)
- Tanis, M. (2008). Health-related on-line forums: What’s the big attraction? *Journal of Health Communication*, 13(7), 698–714. <https://doi.org/10.1080/10810730802415316>
- Torous, J., Jän Myrick, K., Raueo-Ricupero, N., & Firth, J. (2020). Digital mental health and COVID-19: Using technology today to accelerate the curve on access and quality tomorrow. *JMIR Mental Health*, 7(3), e18848. <https://doi.org/10.2196/18848>
- Wagner, B., Schulz, W., & Knaevelsrud, C. (2012). Efficacy of an Internet-based intervention for posttraumatic stress disorder in Iraq: A pilot study. *Psychiatry Research*, 195(1–2), 85–88. <https://doi.org/10.1016/j.psychres.2011.07.026>
- Windler, C., Clair, M., Long, C., Boyle, L., & Radovic, A. (2019). Role of moderators on engagement of adolescents with depression or anxiety in a social media intervention: Content analysis of web-based interactions. *JMIR Mental Health*, 6(9), e13467. <https://doi.org/10.2196/13467>
- Wright, K. (2000). Perceptions of on-line support providers: An examination of perceived homophily, source credibility, communication and social support within on-line support groups. *Communication Quarterly*, 48(1), 44–59. <https://doi.org/10.1080/01463370009385579>
- Wright, K. (2016). Communication in health-related online social support groups/communities: A review of research on predictors of participation, applications of social support theory, and health outcomes. *Review of Communication Research*, 4, 65–87. <https://doi.org/10.12840/issn.2255-4165.2016.04.01.010>
- Wright, K. B., & Miller, C. H. (2010). A measure of weak-tie/strong-tie support network preference. *Communication Monographs*, 77(4), 500–517. <https://doi.org/10.1080/03637751.2010.502538>
- Yew, G. C. K. (2021). Trust in and ethical design of carebots: The case for ethics of care. *International Journal of Social Robotics*, 13(4), 629–645. <https://doi.org/10.1007/s12369-020-00653-w>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.