Enhancing Spoken Communication Experience in Mental Health Peer Support

Manas Satish Bedmutha
UC San Diego
La Jolla, USA
mbedmutha@ucsd.edu

Angelina Van Dyne
UC San Diego Health
La Jolla, USA
avvandyne@health.ucsd.edu

Federica Klaus
UC San Diego Health
La Jolla, USA
fklaus@health.ucsd.edu

Lisa T.Eyler
UC San Diego Health
La Jolla, USA
lteyler@health.ucsd.edu

Nadir Weibel
UC San Diego
La Jolla, CA, USA
weibel@ucsd.edu

Abstract

Peer support is a crucial component of mental health care, offering a more accessible and readily available alternative to professionally-led therapy. Its effectiveness depends not only on verbal exchanges but also on the social communicative behaviors that shape interactions between peer supporters and seekers. While chat-based applications have been widely studied to support peer support in mental health contexts, their inability to capture nonverbal exchanges limits their capacity to fully replicate the depth of in-person conversations. In contrast, spoken communication methods—such as Zoom, face-to-face meetings, and telehealth—are becoming increasingly prevalent in mental health interventions, providing richer and more immersive opportunities for engagement. This shift enables a deeper exploration of communication metrics that incorporate nonverbal cues, offering new ways to assess and enhance peer support interactions. In this paper, we demonstrate how comprehensive modeling of the affective space in peer support conversations allows us to identify key opportunities for improving engagement, fostering emotional connection, and ultimately enhancing the overall mental well-being of participants.

Keywords

Peer Support, Communication Analysis, Non-verbal communication, Social Signals

ACM Reference Format:
Manas Satish Bedmutha, Angelina Van Dyne, Federica Klaus, Lisa T.Eyler, and Nadir Weibel. 2025. Enhancing Spoken Communication Experience in Mental Health Peer Support. In *Proceedings of (CHI '25 Workshop on Envisioning the Future of Interactive Health)*. ACM, New York, NY, USA, 4 pages.

1 Motivation

In the last few years, mental health challenges have been on a constant rise among many populations [7]. While many forms of clinical and subclinical therapies are available, there are various cultural and systemic barriers to access them. Peer support, however, is a form of therapy that is easier to access and more widely available [10]. As digital tools become integral to our life, online peer support on social communities has become prevalent. Face-to-face peer support via online video conferencing has also gained ground with increasing use of telehealth systems.

Peer support therapies, in formalized settings, work on the idea of people helping and supporting others with lived experience of similar illnesses with the help of peer workers who can ensure a safe space to talk about these experiences. Peer support meetings can be conducted 1-1 or in group settings. Trained peer workers are called *facilitators* and the ones seeking help are called *seekers* or *sharers*. When conducted in group settings, each seeker has the opportunity to share while the rest are in an active *listener* role. Among others, empathic support, active engagement and a non-judgmental space contribute immensely to the success of a support session [6]. To achieve success, facilitators often use motivational interviewing, cognitive reframing, and other techniques based on applied cognitive behavioral therapy (CBT) [16].

Although there is a long history of evaluating aspects of professionally led group therapy sessions (e.g., examining therapist-group rapport, group cohesion, or fidelity to a certain therapeutic modality via human review of audio/video tapes), peer support has less often been examined for specific communication features that promote successful outcomes. The ability to computationally model various components of a session can, therefore, help extensively in understanding and empowering participants to achieve their goals at scale.

In online text-based settings, recent advances [13, 14, 16] have allowed modeling of different behaviors exhibited by seekers and facilitators. They have also led to CBT-based solutions for rewriting and reframing one's thoughts in text conversations [14]. While text-based interaction might be preferred by some, others may prefer face-to-face interactions with a peer supporter. Typically spoken-communication, be it in telephonic, video-conferences or in-person settings, provides an added context of non-verbal behaviors, including, for some, how something is being communicated might be the most helpful aspect of the support they receive. In addition, despite some successes, research showed how better results and experiences can be achieved by adding a face or a voice to *bare name* that is usually present in chats [1].

CHI '25 Workshop on Envisioning the Future of Interactive Health, Yokohama, Japan

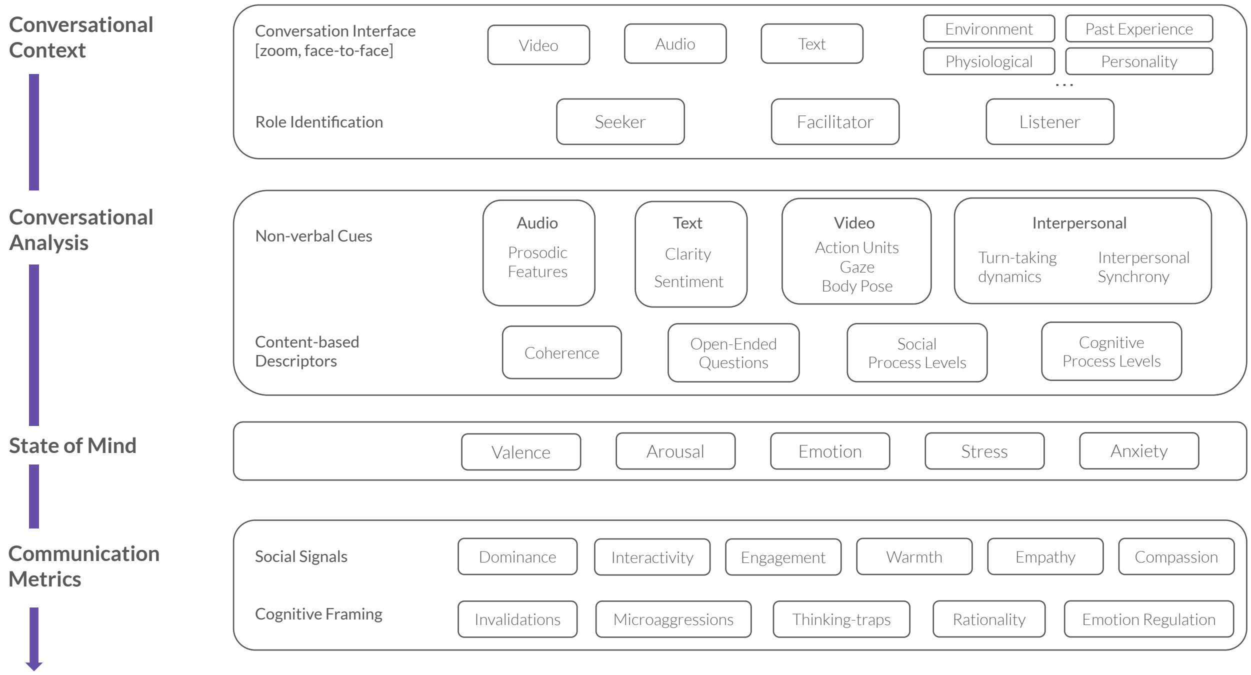

Figure 1: Proposed approach to assess conversational outcomes in peer support settings. The data streams and surrounding context from a conversation can be featurized into verbal and non-verbal cues that can then be used to infer state-of-mind of each individual; collectively these can then be used to model various communication outcomes. These outcomes, or metrics, can be used to influence various interactions in a conversational interface.

Despite the richer context available in face-to-face interaction, the increased difficulty in collecting data in these contexts led to restricting the scale of datasets, and resulted in limited studies and interventions. To the best of our knowledge, there are currently no studies that model spoken communication in peer support groups.

As we gear up to engage with data from such conversational settings 'in the wild', our formative work explores different metrics of communication that can help understand conversational outcomes and provide context to design interactions and help both facilitators and speakers towards improved outcomes. Specifically, we explore what metrics can eventually be computed through a framework leveraging non-verbal cues. These metrics can inspire different interactions for facilitators to promote successful meeting outcomes, for seekers to achieve their goals, and for other peers to become more empathic.

2 Overall Assessment Approach

The key competencies required for effective performance by facilitators or peer-workers are, among others: recovery-oriented, personalized (and person-centered), relationship- and compassion-focused, and trauma-informed [15]. As a result, communicative metrics of success should focus on creating a judgment-free space and helping seekers reframe their process of thinking about their situations. In addition, continuous measures of emotive states of the sharer can help passively inform the efficacy of each individual's state. Building on these principles, we discuss the potential metrics that can be assessed in a peer support group meeting with a broader focus on non-verbal cues that can provide added information beyond text chat-based exchanges. This approach is informed by recent advances in modeling patient-provider communication [3, 4, 18].

We use different degrees of information to predict outcomes from conversations. Our approach can be broken down into four major steps: (1) unpacking conversational context, (2) analyzing it to extract specific features, (3) deriving the state of mind of each speaker, and (4) using the above features to model communication metrics (see Fig. 1).

Conversational Context – To capture the conversational context, we need to first get access to the conversation's data. Depending on the interface chosen for the peer support meeting, data streams from different modalities might be available: typically audio and transcribed text, and potentially video feeds. We might also be able to capture other context about the participants such as their demographics, the physical environment they are in, their personality, past experiences with peer support groups, and potentially physiological data. Since sessions might have multiple sharers, it is important to identify the role of each participant at any given time: facilitators will typically have speaking turns spread throughout a session, while seekers will have a higher turn density at a specific time in the meeting.

Conversation Analysis – For each speaker we extract different descriptors of their communication. Specifically, given their ability to describe affect and mental states [17], we extract prosodic features from speech. We also aim to model clarity of speech and sentiment using transcripts [18]. Further, if available, video information can help us identify facial expressions, body language, and gaze, all of which can help describe various aspects of a person's cognitive state [5]. The availability of video can also shed light on silent listeners in group settings and how their non-verbal cues might impact the perceived relief by the sharer. In addition to exploring the above person-level descriptors, we also aim to capture the interpersonal dynamics that can describe the social relationships in the conversations. These could range from explicit measures related to turn-taking [4] or implicit measures such as synchrony in the non-verbal behaviors. Similarly, we also aim to capture the content of actual spoken information using techniques such as LIWC [11] that can help understand insights into the linguistic style used, the relationships and activities discussed, and the cognitive processes reflected in the language. These elements can inform us about the speakers' reflection process and their fidelity to particular models of peer support [9].

State of Mind – As an intermediary to modeling the communication outcomes themselves, we aim to infer the state of mind for each individual peers in the conversation. State of mind could be determined by the valence, arousal and emotional changes of the participants [12], or their perceived stress and anxiety during the peer support process [8]. Inferring a specific state of mind can help both support seekers and facilitators. Support seekers might be able to identify what enables them to achieve their goals and help them open up – their peers' behaviors, specific style of facilitation (motivational interviewing, active listening, etc.), or their own body language; if they track their state of mind changes over time, they can also investigate their progress, resilience, and regulation abilities. Facilitators can use the inferred state of mind of support seekers as a key indicator of their performance, such as if they are making progress and identifying causal indicators of their behaviors that lead to improved conversation outcomes.

Communication Metrics – We aim to identify two forms of communication metrics. The first group relates to the emotive behaviors exhibited during social settings, or *social signals*; these include person-centered behaviors such as dominance, interactivity, engagement, warmth [4] or interpersonal dynamics such as empathy and compassion [13]. The second type of metrics involves the CBT-inspired concept of cognitive reframing that aims to reduce falling into thinking-traps or being rational and emotionally regulated about their actions [14]. Additionally, to investigate wether the meeting created an open space, it is important to identify the presence of microaggressions or invalidations [3] .

The described 4-steps process is highly non-linear, meaning that non-verbal cues might influence the state of mind as well as communication metrics in different ways. Our proposed approach highlighted in Figure 1 should therefore be considered a starting point to explore the interplay of affect and predictors: each cognitive behavior could be used as a latent descriptor to identify new causal relationships between affect, and emotion, and how they present in our body language. Additionally, given how our framework is based on multiple layers of abstraction, each of them can be used as context to inform different interactions with our end-users (sharers and facilitators).

3 Next Steps: Exploring affect-aware interactions in the real-world

We described our approach to enhance spoken communication experience in mental health support groups by presenting a model that extracts from each peer group session non-verbal cues, state of mind and communication metrics, and how these can inform the efficacy of treatment, facilitation, and success strategies to improve peer support experiences. We believe that (i) each of these elements can inform critical feedback for different roles to achieve their goals better, and (ii) such interactions can be visualized in real-time, after a session or over-time [2] to help both facilitators and support seekers.

In our work, we are investigating how to leverage existing models for computing various affects while fine-tuning our approach to a specific dataset of online peer support meetings collected by Recovery International.[1] We are focusing on how to represent this information in real-time by using specific visual nudges that help participants to be mindful of interruptions, or introducing subtle guidance about cognitive reframing [14]. We are also exploring tracking the temporal progress of each participant over a session and use that information to inform personalized facilitation.

References

[1] A. Banbury, L. Parkinson, S. Gordon, and D. Wood. Implementing a peer-support programme by group videoconferencing for isolated carers of people with dementia. *Journal of telemedicine and telecare*, 25(9):572–577, 2019.
[2] M. S. Bedmutha, E. Bascom, C.-P. Reggie, K. Tobar, J. Sabin, W. Pratt, B. Wood, A. Hartzler, and N. Weibel. Towards designing visualizations to understand social signals in patient-provider communication. 2023.
[3] M. S. Bedmutha, E. Bascom, K. R. Sladek, K. Tobar, R. Casanova-Perez, A. Andreiu, A. Bhat, S. Mangal, B. R. Wood, J. Sabin, et al. Artificial intelligence-generated feedback on social signals in patient–provider communication: technical performance, feedback usability, and impact. *JAMIA open*, 7(4):ooae106, 2024.
[4] M. S. Bedmutha, A. Tsedenbal, K. Tobar, S. Borsotto, K. R. Sladek, D. Singh, R. Casanova-Perez, E. Bascom, B. Wood, J. Sabin, et al. Conversense: An automated approach to assess patient-provider interactions using social signals. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–22, 2024.
[5] M. Cristani, R. Raghavendra, A. Del Bue, and V. Murino. Human behavior analysis in video surveillance: A social signal processing perspective. *Neurocomputing*, 100:86–97, 2013.
[6] D. Deng, T. Rogers, and J. A. Naslund. The role of moderators in facilitating and encouraging peer-to-peer support in an online mental health community: a qualitative exploratory study. *Journal of Technology in Behavioral Science*, 8(2):128–139, 2023.
[7] I. Ferwana and L. R. Varshney. The impact of covid-19 lockdowns on mental health patient populations in the united states. *Scientific reports*, 14(1):5689, 2024.
[8] A. Irajpour, M. S. Hashemi, P. Abazari, S. Shahidi, and M. Fayazi. The effects of peer support on depression, anxiety, and stress among patients receiving hemodialysis. *Iran Red Crescent Med J*, 20(S1):e66321, 2018.
[9] A. Konopasky, S. J. Durning, A. R. Artino, D. Ramani, and A. Battista. The linguistic effects of context specificity: exploring affect, cognitive processing, and agency in physicians' think-aloud reflections. *Diagnosis*, 7(3):273–280, 2020.
[10] R. Merchant, A. Goldin, D. Manjanatha, C. Harter, J. Chandler, A. Lipp, T. Nguyen, and J. A. Naslund. Opportunities to expand access to mental health services: A case for the role of online peer support communities. *Psychiatric Quarterly*, 93(2):613–625, 2022.
[11] J. W. Pennebaker, M. E. Francis, and R. J. Booth. Linguistic Inquiry and Word Count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.
[12] R. S. Sahi, N. I. Eisenberger, and J. A. Silvers. Peer facilitation of emotion regulation in adolescence. *Dev. Cogn. Neurosci.*, 62:101262, 2023.
[13] A. Sharma, A. S. Miner, D. C. Atkins, and T. Althoff. A computational approach to understanding empathy expressed in text-based mental health support. *arXiv preprint arXiv:2009.08441*, 2020.
[14] A. Sharma, K. Rushton, I. W. Lin, D. Wadden, K. G. Lucas, A. S. Miner, T. Nguyen, and T. Althoff. Cognitive reframing of negative thoughts through human-language model interaction. *arXiv preprint arXiv:2305.02466*, 2023.

[1] https://recoveryinternational.org/

[15] Substance Abuse and Mental Health Services Administration. Core competencies for peer workers in behavioral health services, 2018.

[16] S. Syed, Z. Iftikhar, A. W. Xiao, and J. Huang. Machine and human understanding of empathy in online peer support: A cognitive behavioral approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2024.

[17] M. Valstar, B. Schuller, K. Smith, F. Eyben, B. Jiang, S. Bilakhia, S. Schnieder, R. Cowie, and M. Pantic. Avec 2013: the continuous audio/visual emotion and depression recognition challenge. In *Proc. of the 3rd ACM international workshop on Audio/visual emotion challenge*, pages 3–10, 2013.

[18] Z. Wang, N. Hassan, V. LeBaron, T. Flickinger, D. Ling, J. Edwards, C. Wu, M. Boukhechba, and L. E. Barnes. Commsense: A wearable sensing computational framework for evaluating patient-clinician interactions. *Proc. of the ACM on Human-Computer Interaction*, 8(CSCW2):1–31, 2024.