Human-AI Collaboration Enables More Empathic Conversations in Peer-to-Peer Mental Health Support

Tim Althoff

@timalthoff

University of Washington

behavioral data science

NASEM Forum on Mental Health and Substance Use Disorders, April 2022

Acknowledgements

Ashish Sharma

Inna Lin

Adam Miner

Dave Atkins

TalkLife

behavioral data science

UW Medicine

BRiTE

Microsoft Research

Bill & Melinda Gates foundation

NIH

W UNIVERSITY of WASHINGTON

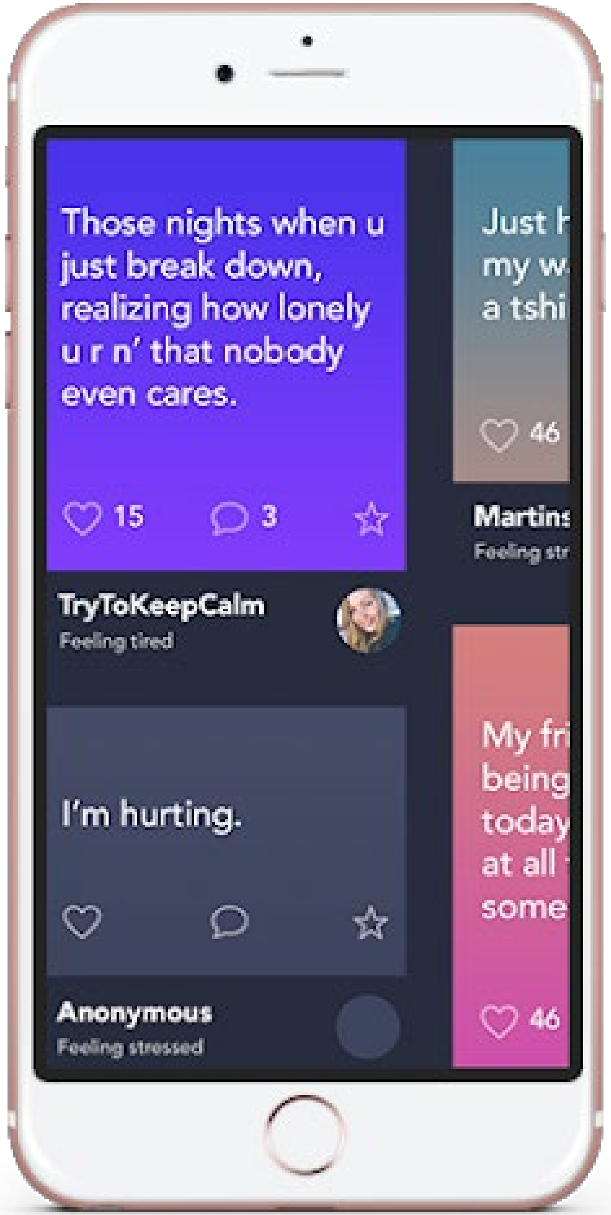
Ai2

Content Warning

This talk contains anonymized examples related to mental illness, self-harm and suicidal ideation.

Mental Health: Need vs. Access

- **Access** to **mental health care** is **poor** across the globe
  - We may **never** have enough mental health professionals to meet the **increasing need**
- **Online peer support platforms can help!**
  - Millions of peers seek and provide support through conversations



Those nights when u just break down, realizing how lonely u r n' that nobody even cares.

♡ 15 💬 3 ☆

TryToKeepCalm
Feeling tired

I'm hurting.

♡ 💬 ☆

Anonymous
Feeling stressed

Just h
my w
a tshi

♡ 46

Martins
Feeling str

My fri
being
today
at all
some

♡ 46

Key Motivation

- Peer supporters on these platforms are amazing and volunteer their time and energy to support others in need.

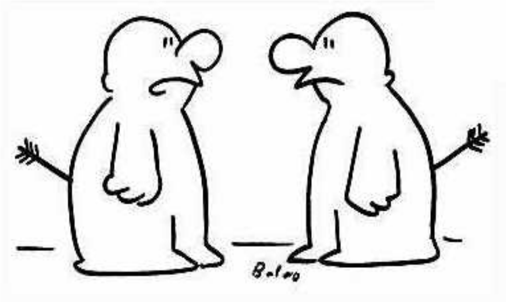
- However, they are rarely trained to provide *effective* support.

- Could technology help supporters support others better?

- How could we turn connection into meaningful and effective interactions?

Empathy

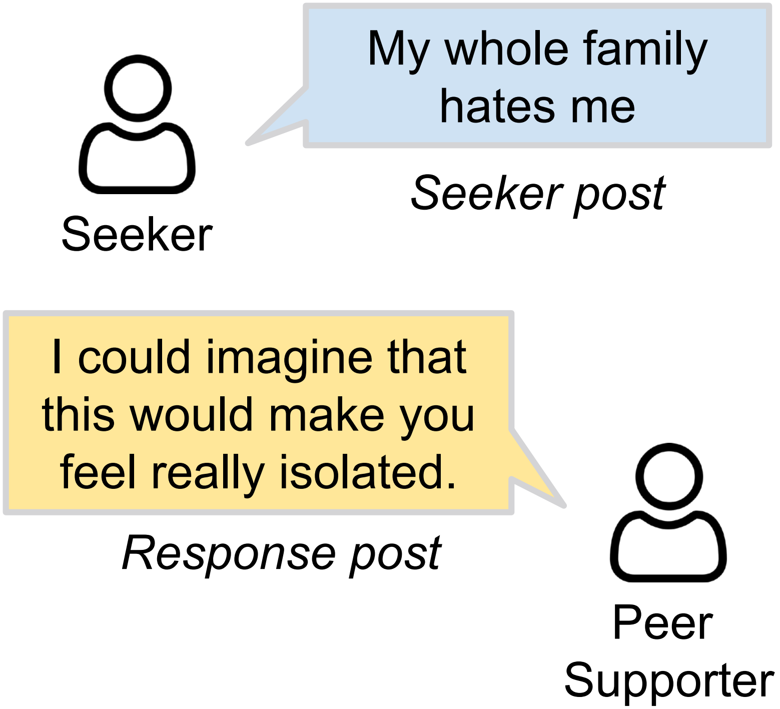
Empathy: The ability to **understand** or **feel** the emotions and experiences of others



"I know exactly how you feel."

High empathy interactions

- ○ Strong associations with **positive counseling outcomes** like *alliance* and *rapport* (Bohart et al., 2002; Elliot et al., 2011)



Empathic interaction

(1) How can we measure expressed empathy?

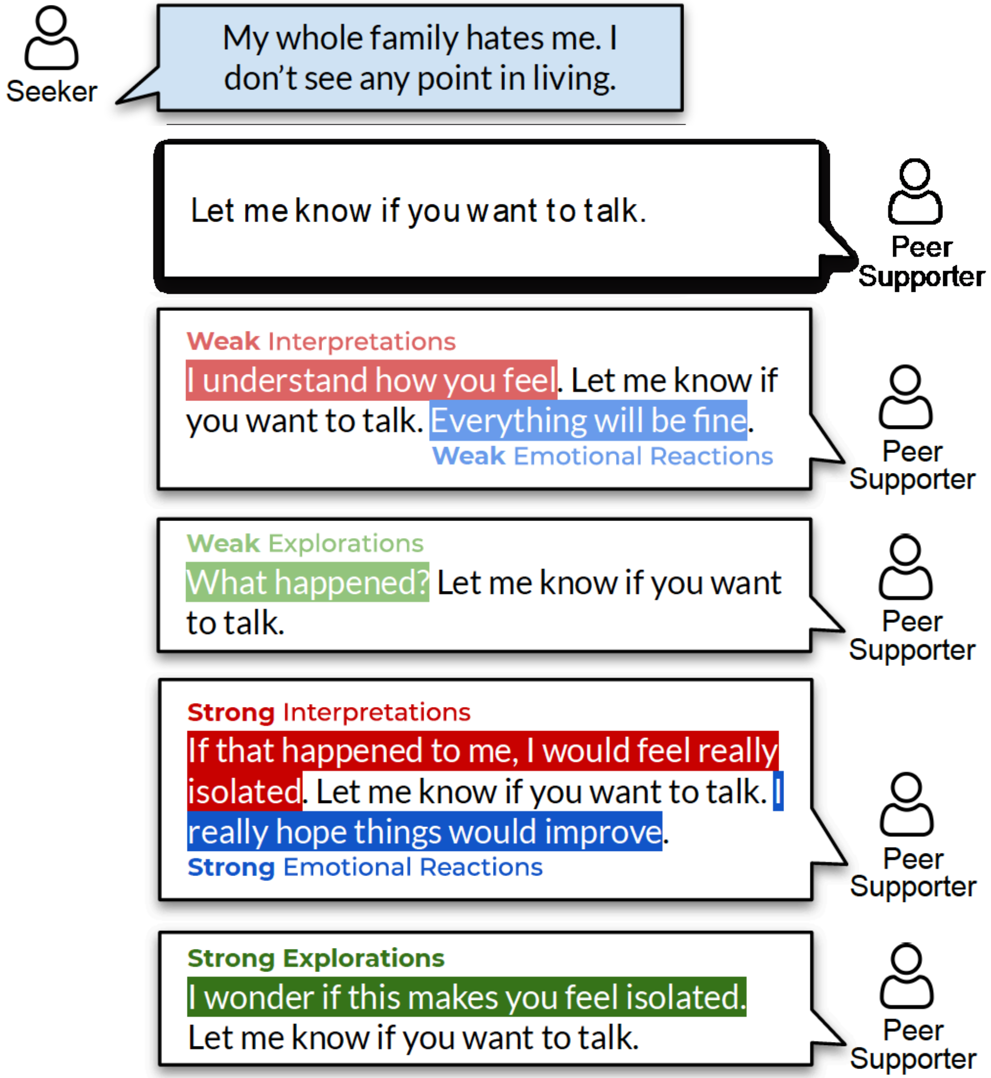
Framework of empathy expressed in conversations

Three communication mechanisms of empathy

- **Emotional Reactions:** communicating the emotions experienced after reading a post
- **Interpretations:** communicating understanding of the inferred feelings / experiences
- **Explorations:** improving one's understanding by exploring feelings / experiences

We differentiate between

- peers **not** expressing them at all (level 0)
- peers expressing them to some **weak** degree (level 1)
- peers expressing them **strongly** (level 2)



My whole family hates me. I don't see any point in living.

Let me know if you want to talk.

Weak Interpretations

I understand how you feel. Let me know if you want to talk. Everything will be fine.

Weak Emotional Reactions

Weak Explorations

What happened? Let me know if you want to talk.

Strong Interpretations

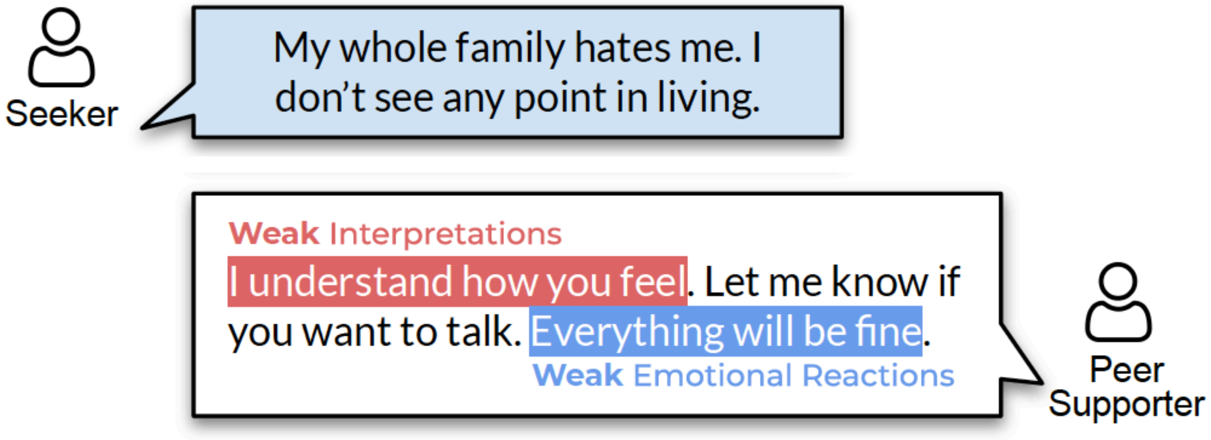
If that happened to me, I would feel really isolated. Let me know if you want to talk. I really hope things would improve.

Strong Emotional Reactions

Strong Explorations

I wonder if this makes you feel isolated. Let me know if you want to talk.

Prediction Tasks



My whole family hates me. I don’t see any point in living.

I understand how you feel. Let me know if you want to talk. Everything will be fine.

Task 1: Empathy Identification

How empathic is response post in the context of seeker post?

Emotional Reactions – 1 out of 2
Interpretations – 1 out of 2
Explorations – 0 out of 2

Task 2: Rationale Extraction

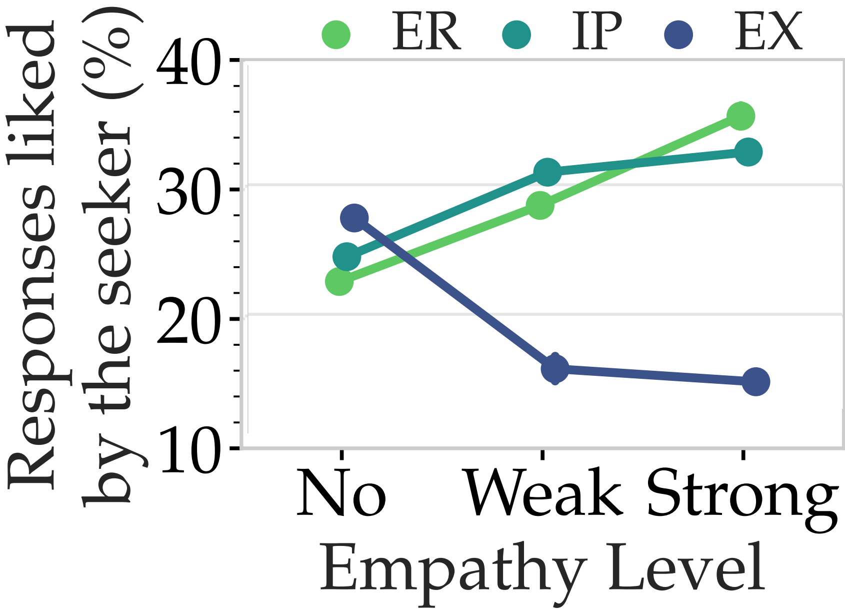
What is the supporting evidence (*rationale*) for the identified empathy levels?

A Computational Approach to Understanding Empathy Expressed in Text-Based Mental Health Support.
Sharma, Miner, Atkins, Althoff. EMNLP 2020.

(2) How is empathy expressed currently on the Talklife platform and what are associated outcomes?

Model-based Insights into Mental Health Platforms

Good News: Empathy appears meaningful to TalkLife users



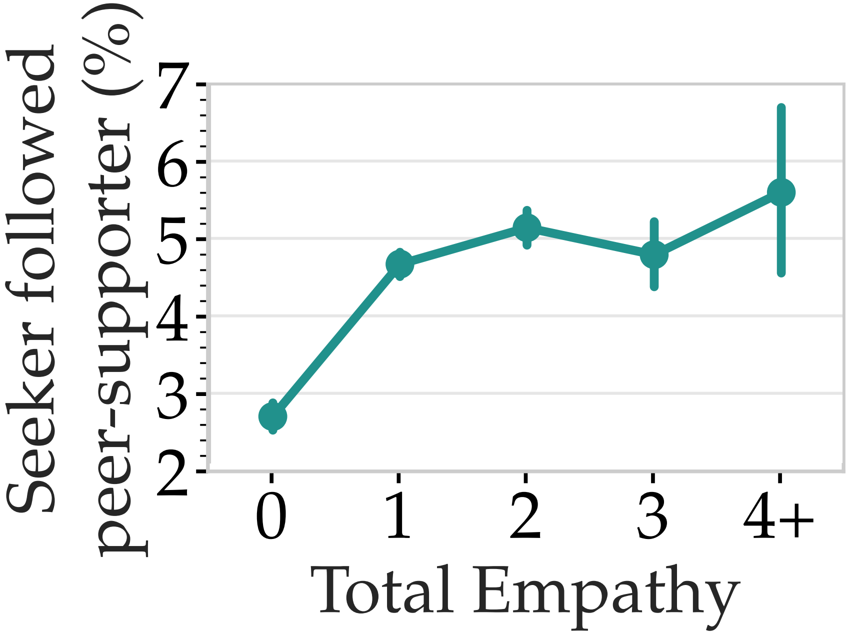
Strong communications of emotional reactions and interpretations receive **45% more likes** than their no communication

Stronger explorations get **47% more replies**

High empathy interactions are received positively by seekers. They drive *engagement* on social media platforms.

Model-based Insights into Mental Health Platforms

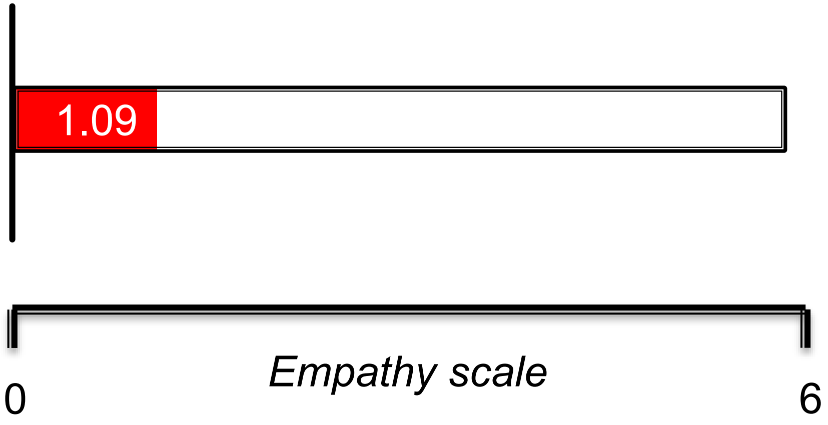
Good News: Empathy appears meaningful to TalkLife users



Seekers are 79% more likely to “follow” peer supporters after an empathic interaction than after a non-empathic one

Relationship forming more likely after empathic interactions

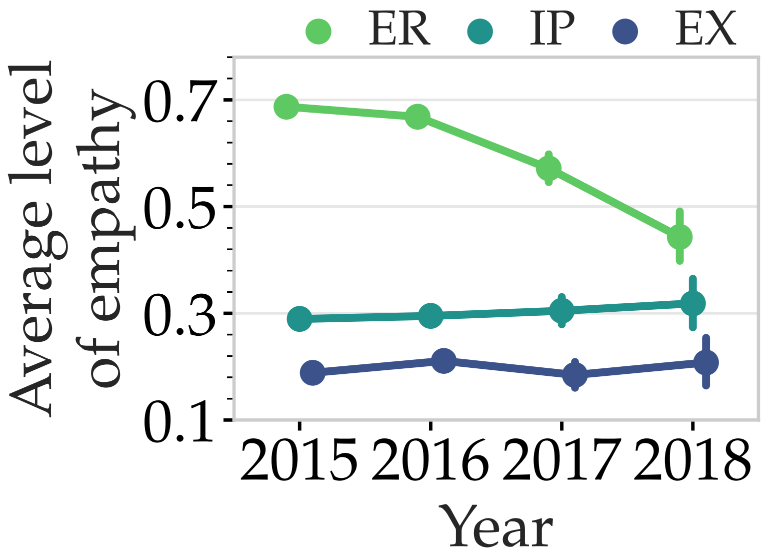
Need for Empathic Feedback and Training



Expressed empathy is typically low

Does it improve over time?

Peer-supporters do not self-learn empathy over time



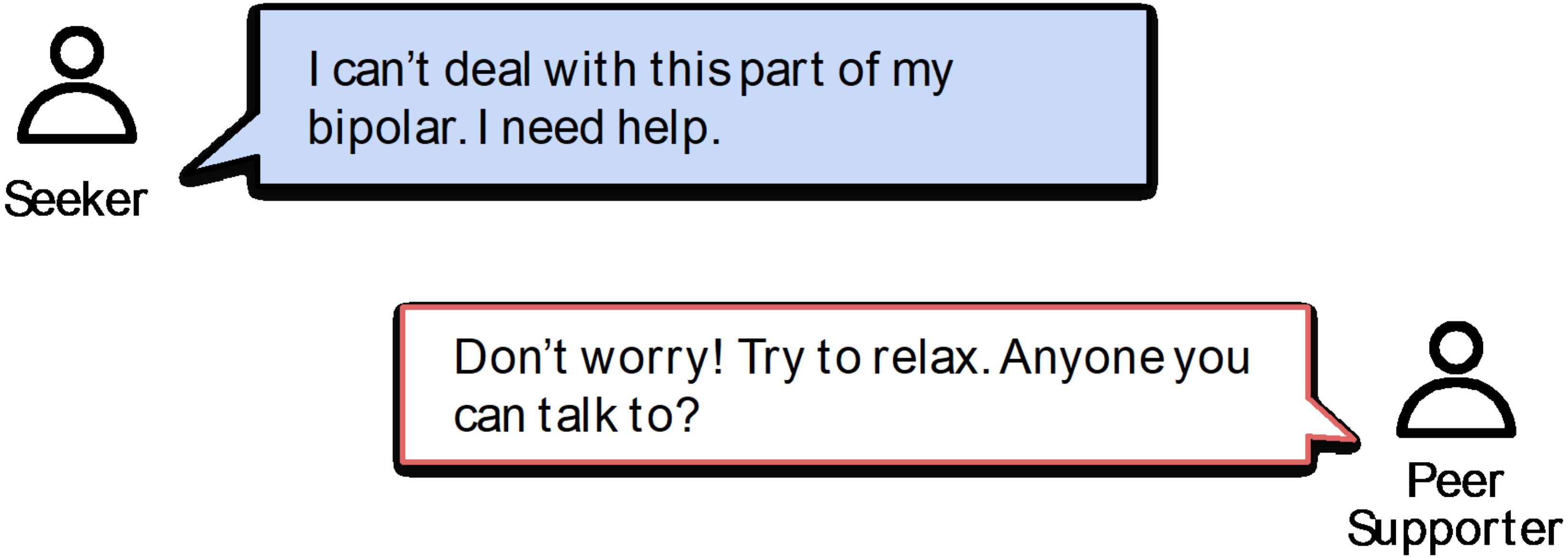
This is also true for therapists!

- Without **deliberate practice** and **specific feedback**, even trained therapists often diminish in skills over time (Goldberg et al., 2016)

(2) How could a machine give feedback on empathy expression?

New Task: Empathic Rewriting

Empathic Rewriting: Computationally transform low-empathy conversational posts to higher empathy

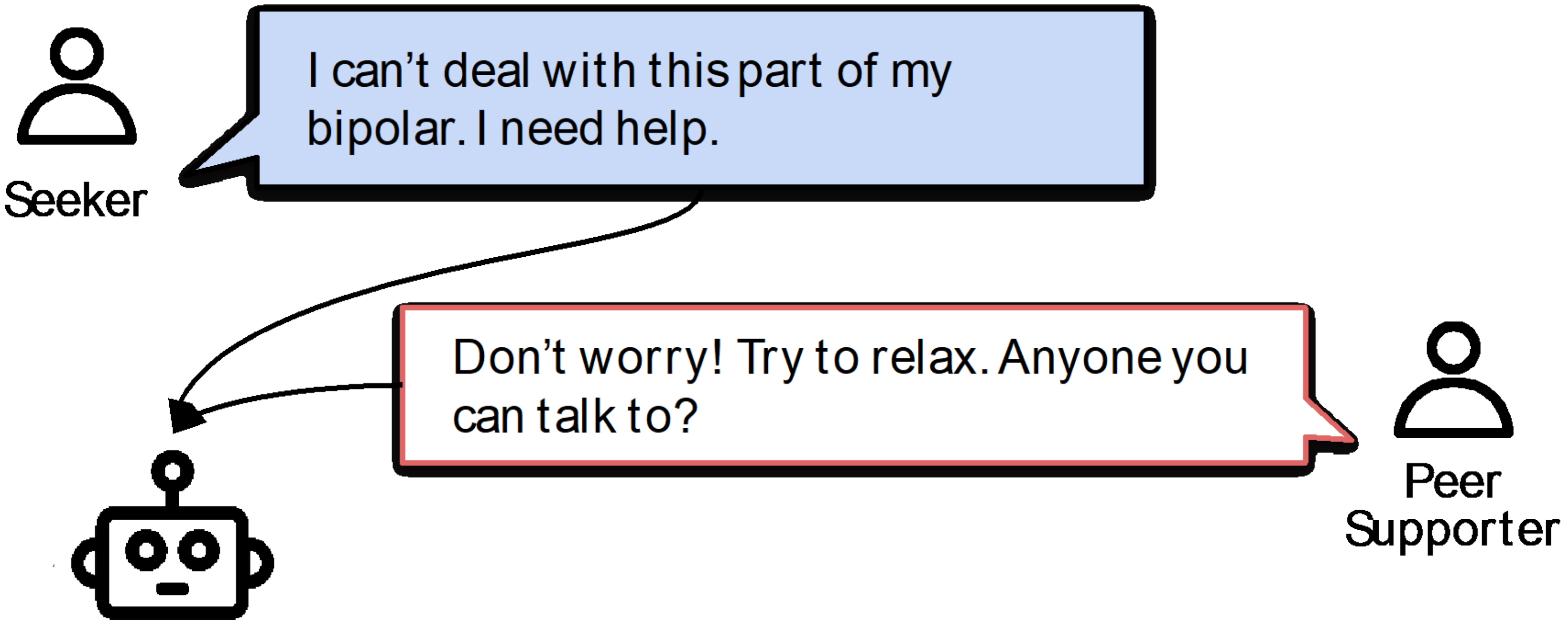


I can’t deal with this part of my bipolar. I need help.

Don’t worry! Try to relax. Anyone you can talk to?

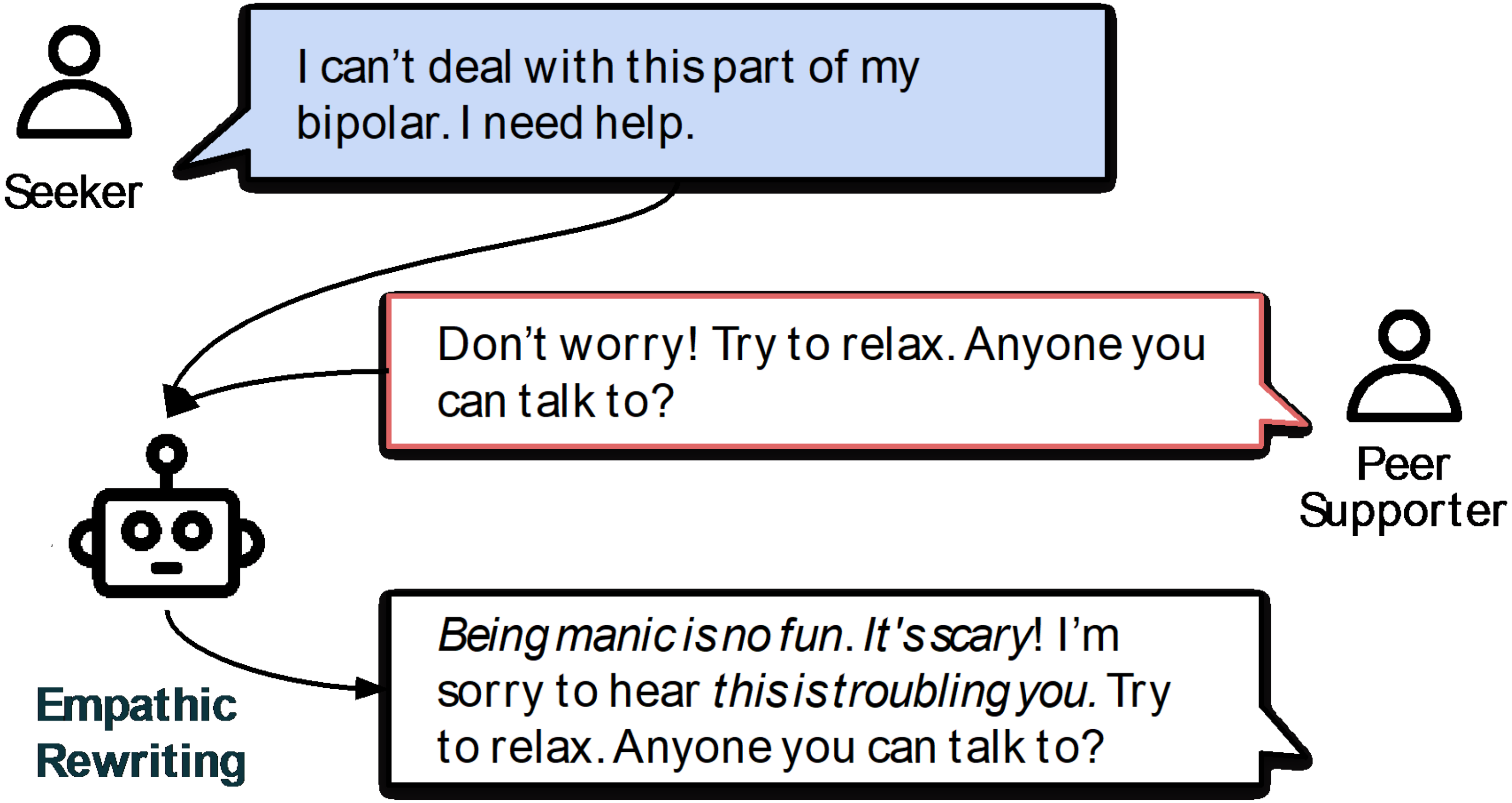
New Task: Empathic Rewriting

Empathic Rewriting: Computationally transform low-empathy conversational posts to higher empathy


Seeker
I can't deal with this part of my bipolar. I need help.
Don't worry! Try to relax. Anyone you can talk to?
Peer Supporter

New Task: Empathic Rewriting

Empathic Rewriting: Computationally transform low-empathy conversational posts to higher empathy



Seeker

I can’t deal with this part of my bipolar. I need help.

Don’t worry! Try to relax. Anyone you can talk to?

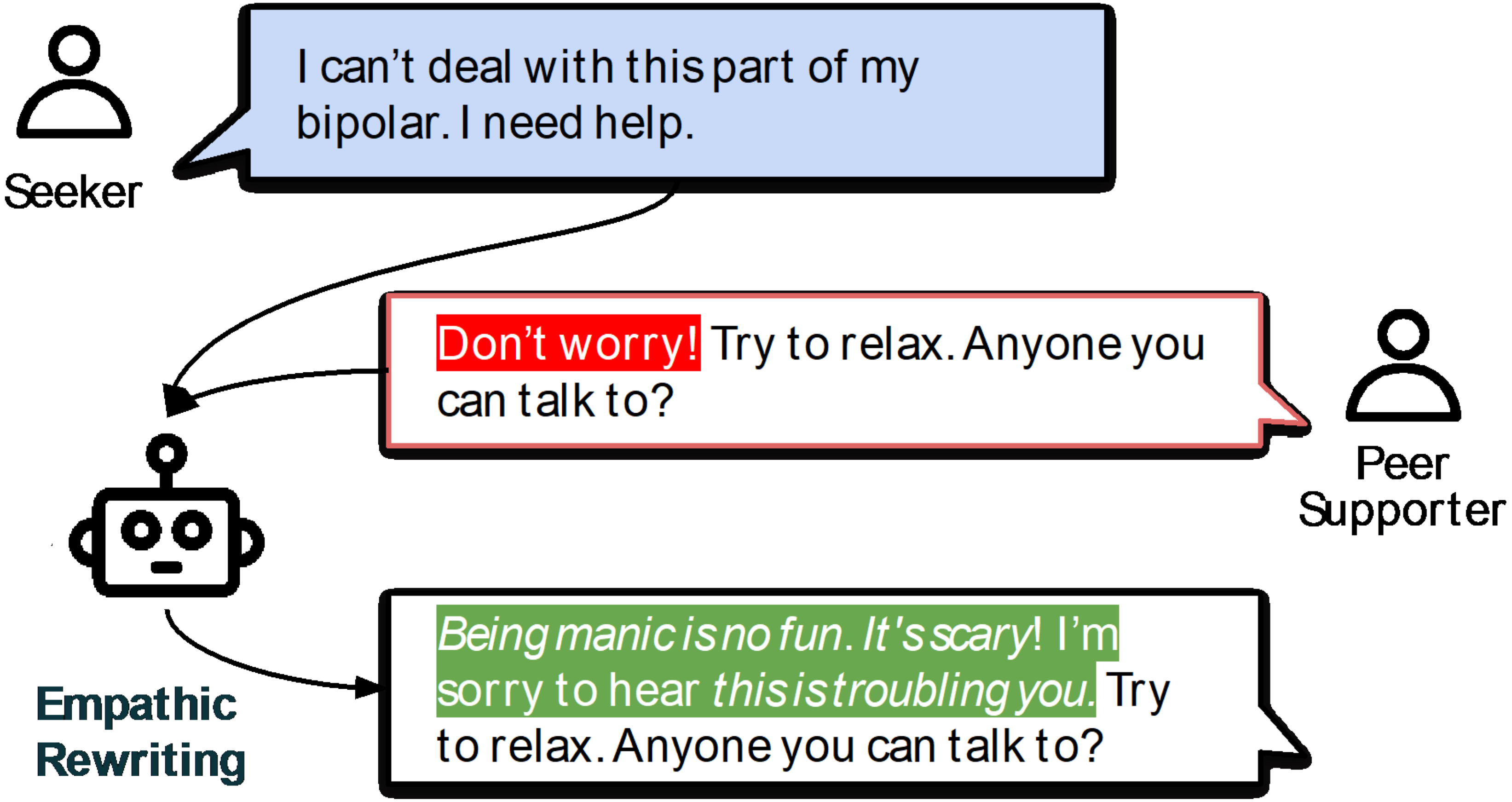
Peer Supporter

Empathic Rewriting

Being manic is no fun. It's scary! I’m sorry to hear *this is troubling you.* Try to relax. Anyone you can talk to?

New Task: **Empathic Rewriting**

Empathic Rewriting: Computationally transform low-empathy conversational posts to higher empathy



I can't deal with this part of my bipolar. I need help.

Don't worry! Try to relax. Anyone you can talk to?

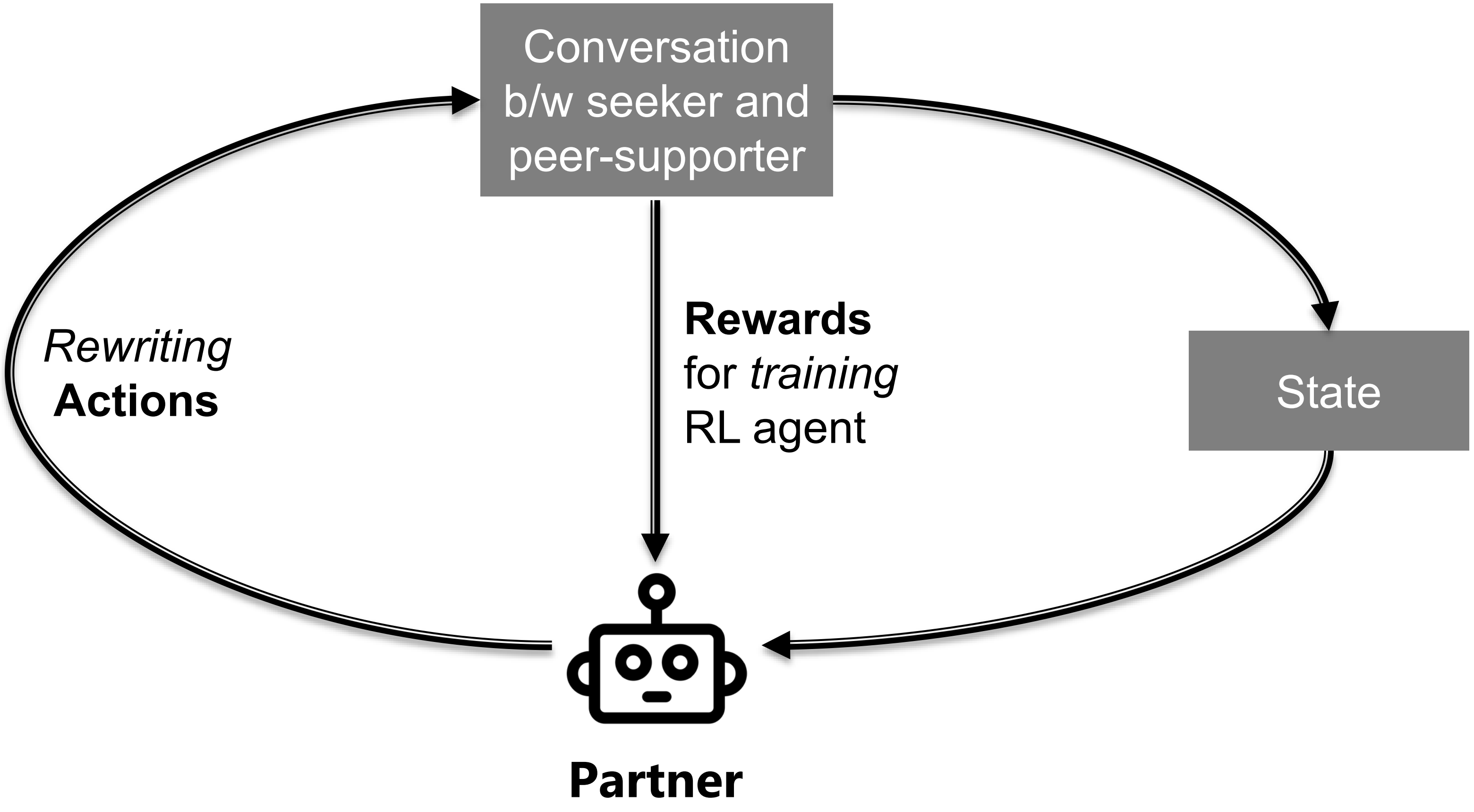
Being manic is no fun. It's scary! I'm sorry to hear *this is troubling you.* Try to relax. Anyone you can talk to?

Text inserted

Text removed

Partner: Empathic Rewriting using Reinforcement Learning (RL)

PARTNER is an **RL agent** for the task of empathic rewriting


Conversation b/w seeker and peer-supporter
Rewriting **Actions**
Rewards for *training* RL agent
State
Partner

PARTNER is **rewarded** for

- Increased **empathy**
- **Fluent** English
- A **coherent** response
- Being **context-specific** (instead of generic responses)

Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. Sharma, Lin, Miner, Atkins, Althoff. WebConf 2021. Best Paper Award.

How to train this model? Where does the data come from?

Increasing empathy of a conversation is **challenging**

Can we do the **reverse process** of **decreasing empathy** instead?

I'm sorry to hear this is troubling you.
Have you…

Original response

~~I'm sorry to hear this is troubling you.~~
Have you…

Empathic sentence removed

Now we can create a **parallel dataset of millions of post-response pairs!**

Based on idea by West & Horvitz (2019)

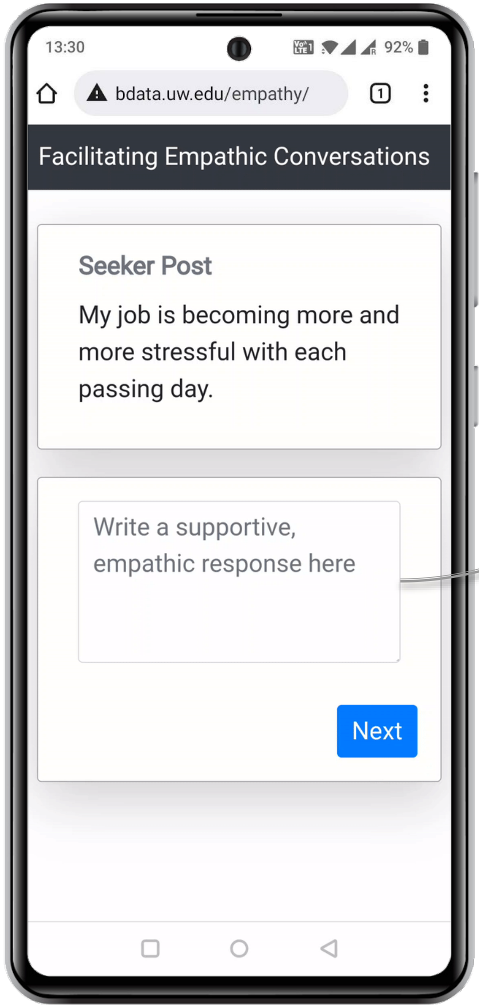
(3) How can peer supporters and AI collaborate?

Key Design Considerations

- How can we keep the meaningful human-human conversation at the center?
- This is not “human-in-the-loop”: It’s AI-in-the-loop, on a back-seat, with human supervision
- How do we give minimal feedback that is maximally effective?
- How do we transparently communicate to potential users around potential benefits, data use, mechanisms to opt in/out, and report concerns?

Peer supporters may **express higher empathy** with **AI-based feedback** (1)

Human Only (Control Group)



Facilitating Empathic Conversations

Seeker Post

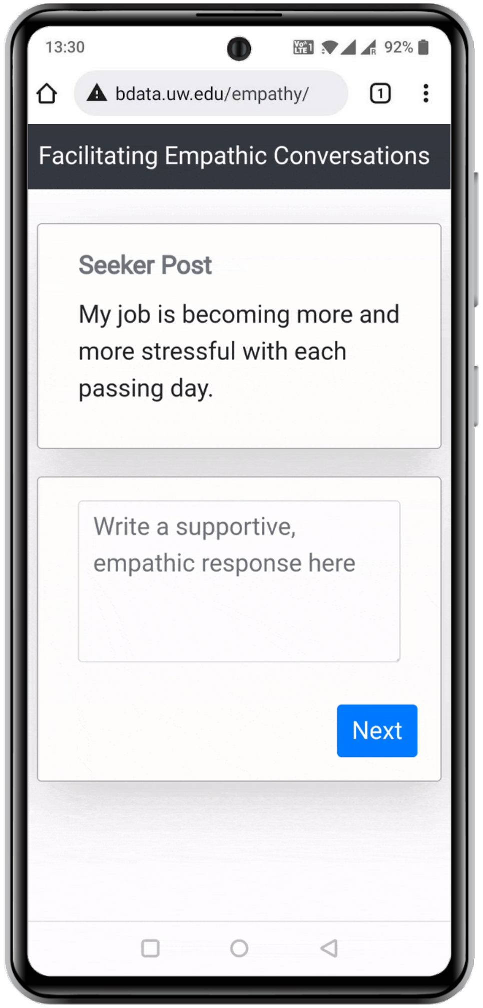
My job is becoming more and more stressful with each passing day.

Write a supportive, empathic response here

Empty Chatbox for Writing Responses

Peer supporters may **express higher empathy** with **AI-based feedback** (2)

Human Only
(Control Group)



13:30 92%

bdata.uw.edu/empathy/

Facilitating Empathic Conversations

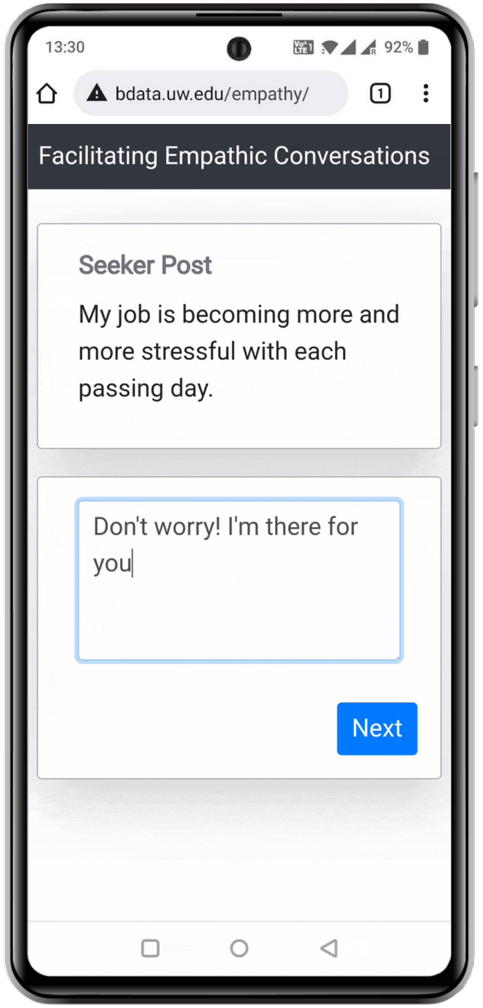
Seeker Post

My job is becoming more and more stressful with each passing day.

Write a supportive, empathic response here

Peer supporters may **express higher empathy** with **AI-based feedback** (3)

Human Only (Control Group)



13:30 92%

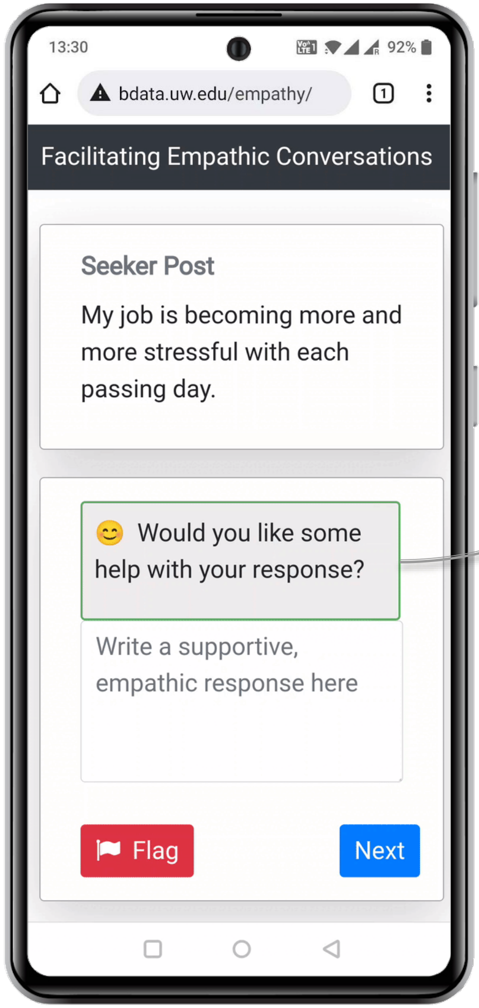
bdata.uw.edu/empathy/

Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

Don't worry! I'm there for you



13:30 92%

bdata.uw.edu/empathy/

Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

😊 Would you like some help with your response?

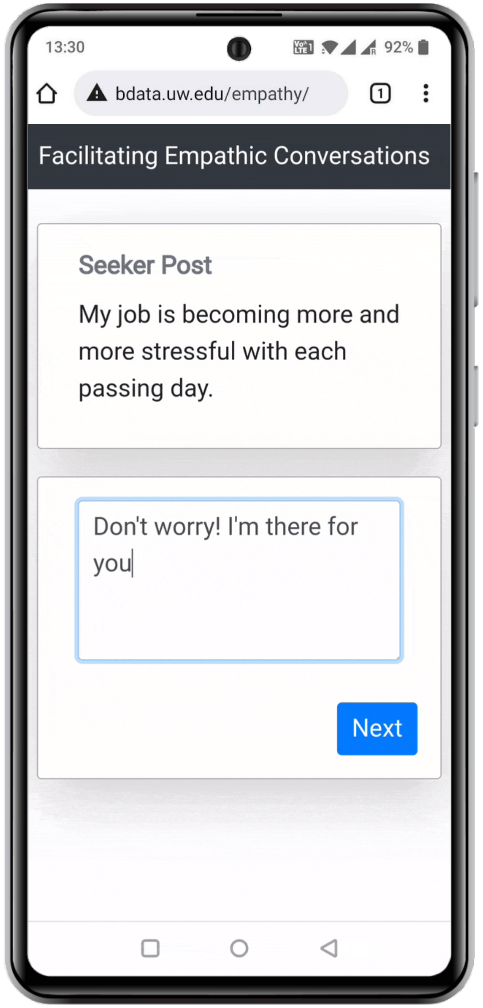
Write a supportive, empathic response here

Human + AI (Treatment Group)

Feedback Prompts

Peer supporters may **express higher empathy** with **AI-based feedback** (4)

Human Only
(Control Group)

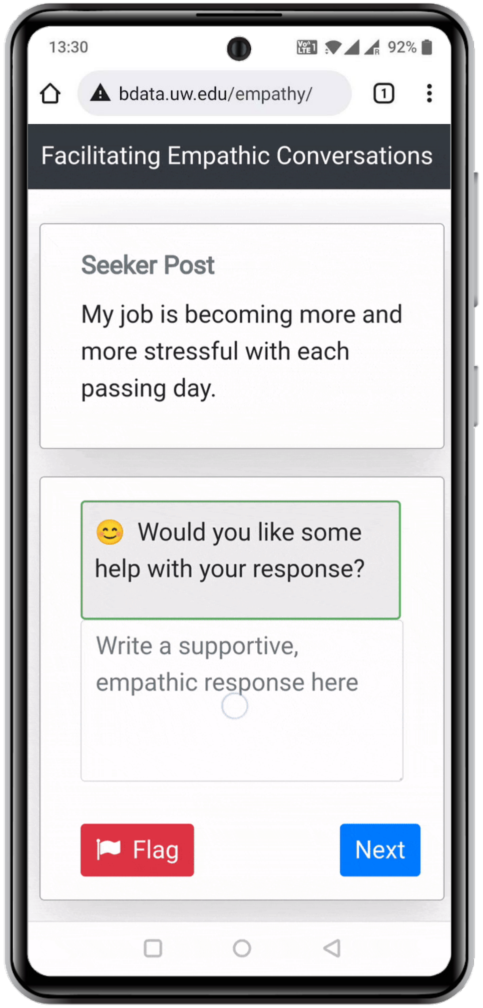


Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

Don't worry! I'm there for you



Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

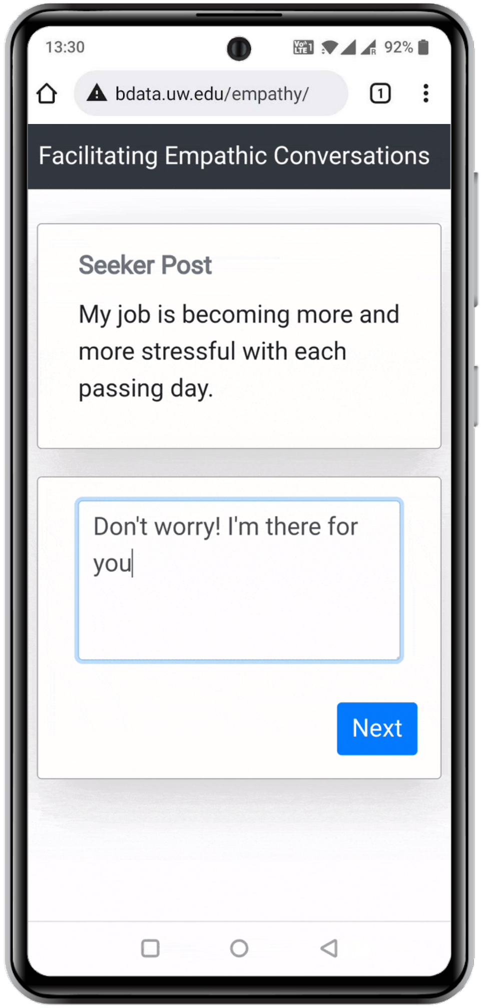
😊 Would you like some help with your response?

Write a supportive, empathic response here

Human + AI
(Treatment Group)

Peer supporters may **express higher empathy** with **AI-based feedback** (5)

Human Only (Control Group)



13:30 92%

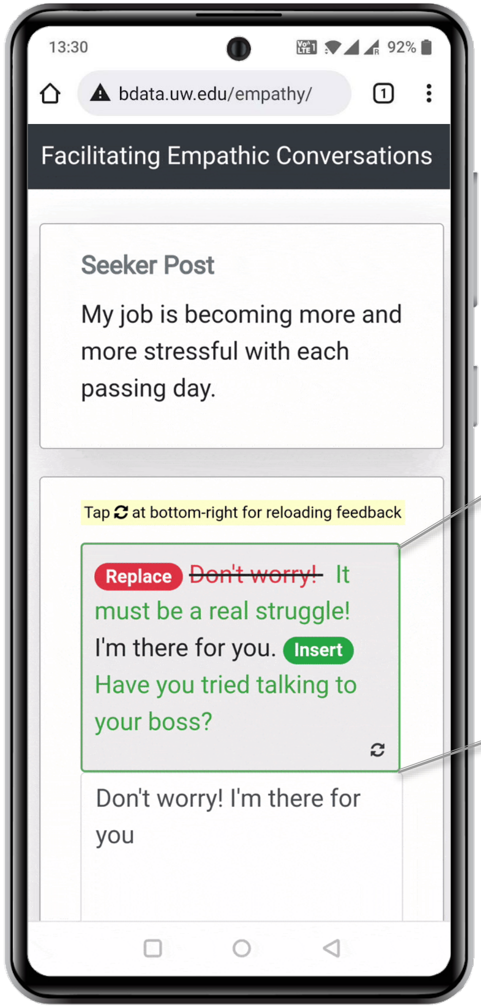
bdata.uw.edu/empathy/

Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

Don't worry! I'm there for you



13:30 92%

bdata.uw.edu/empathy/

Facilitating Empathic Conversations

Seeker Post

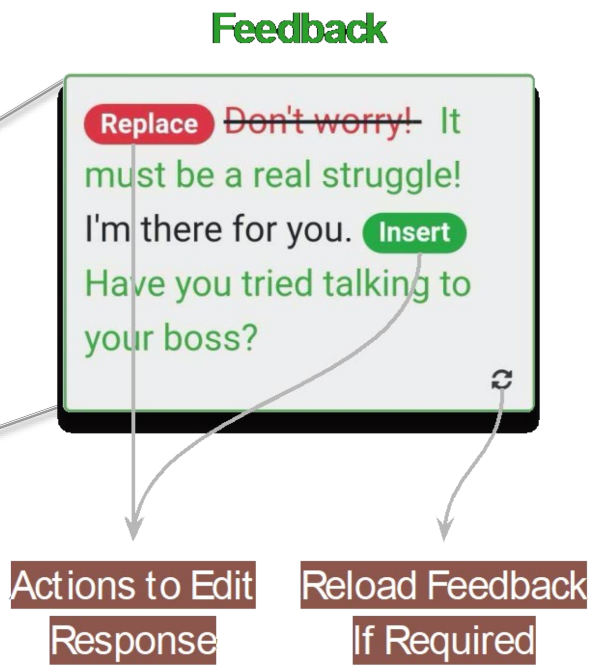
My job is becoming more and more stressful with each passing day.

Tap ⟳ at bottom-right for reloading feedback

Replace ~~Don't worry!~~ It must be a real struggle! I'm there for you. Insert Have you tried talking to your boss?

Don't worry! I'm there for you

Human + AI (Treatment Group)

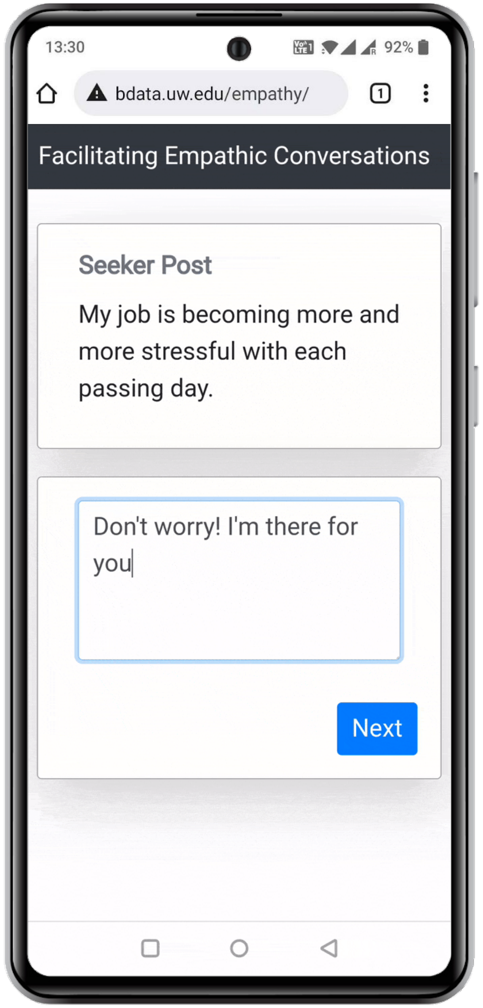


Feedback

Replace ~~Don't worry!~~ It must be a real struggle! I'm there for you. Insert Have you tried talking to your boss?

Peer supporters may **express higher empathy** with **AI-based feedback** (6)

Human Only
(Control Group)

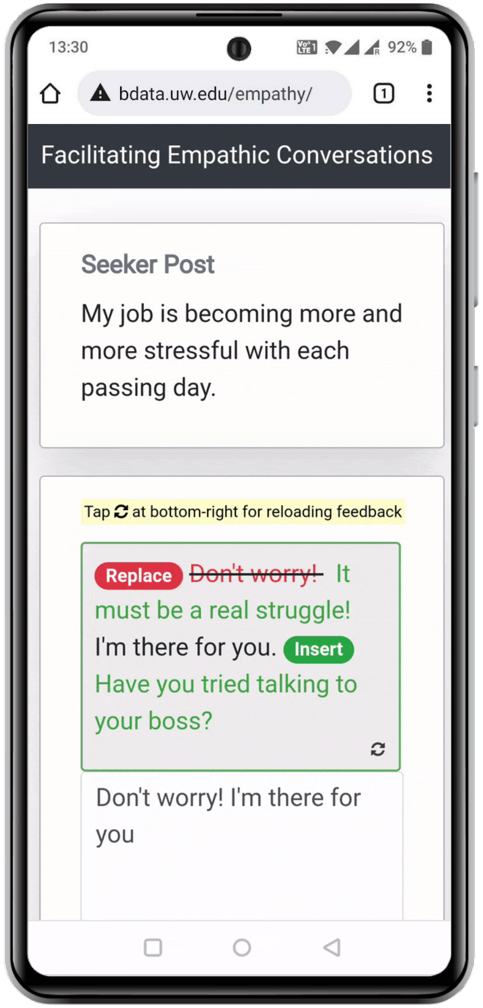


Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

Don't worry! I'm there for you



Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

Tap at bottom-right for reloading feedback

Replace ~~Don't worry!~~ It must be a real struggle! I'm there for you. Insert Have you tried talking to your boss?

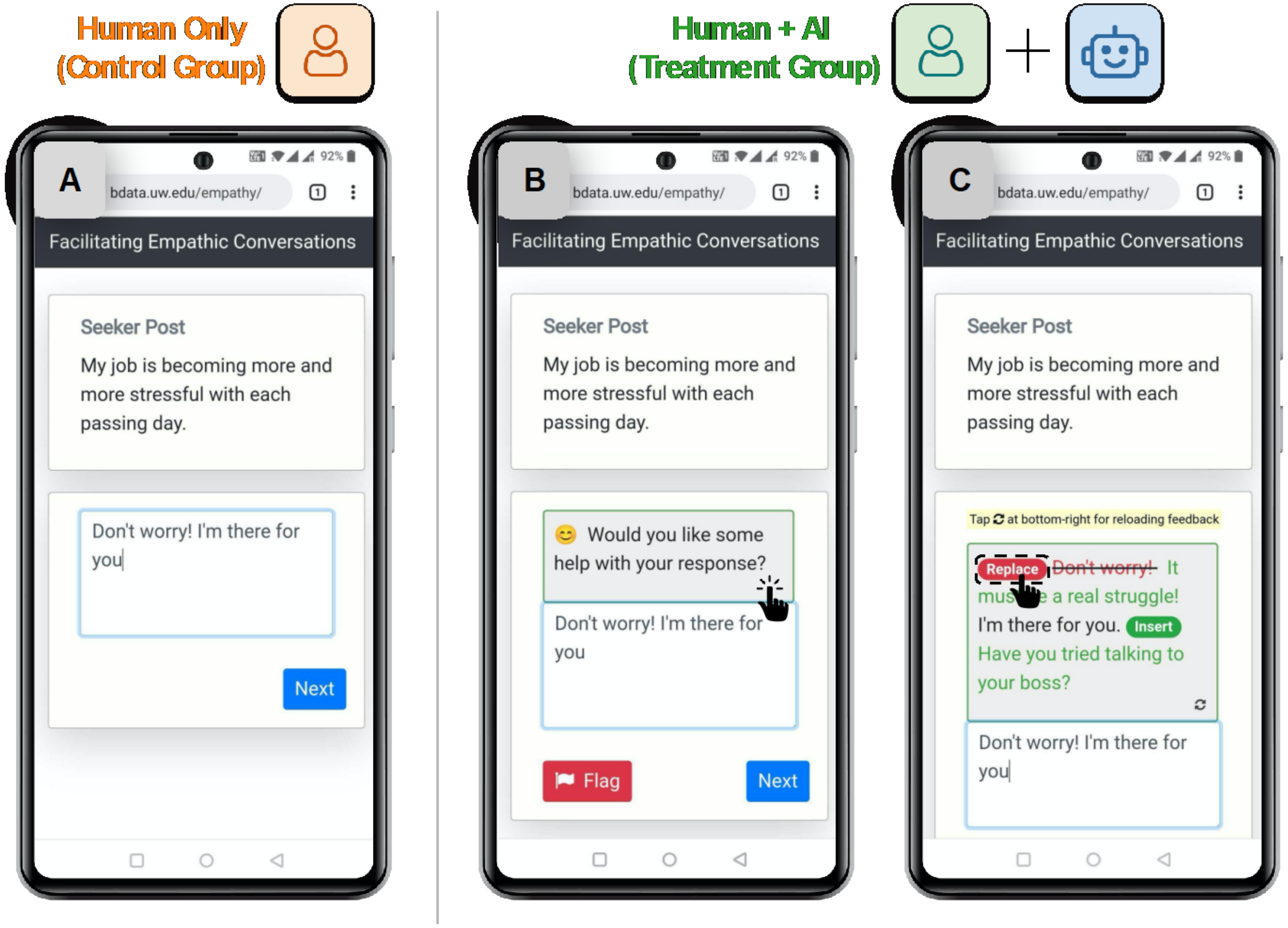
Don't worry! I'm there for you

Human + AI
(Treatment Group)

(4) Does it work?

Study Design: Randomized Controlled Trial for Examining the **Effects of AI-based Feedback on Empathy**

- Recruit participants from TalkLife and randomly divide them into control and treatment groups (N=300)
- Importantly, both groups received empathy training at the beginning.
  - Do concrete real-time suggestions help beyond traditional training methods?
- Participants write responses to 10 existing seeker posts
  - Different posts for different participants
  - Same posts across control and treatment



Facilitating Empathic Conversations

Seeker Post

My job is becoming more and more stressful with each passing day.

Don't worry! I'm there for you

😊 Would you like some help with your response?

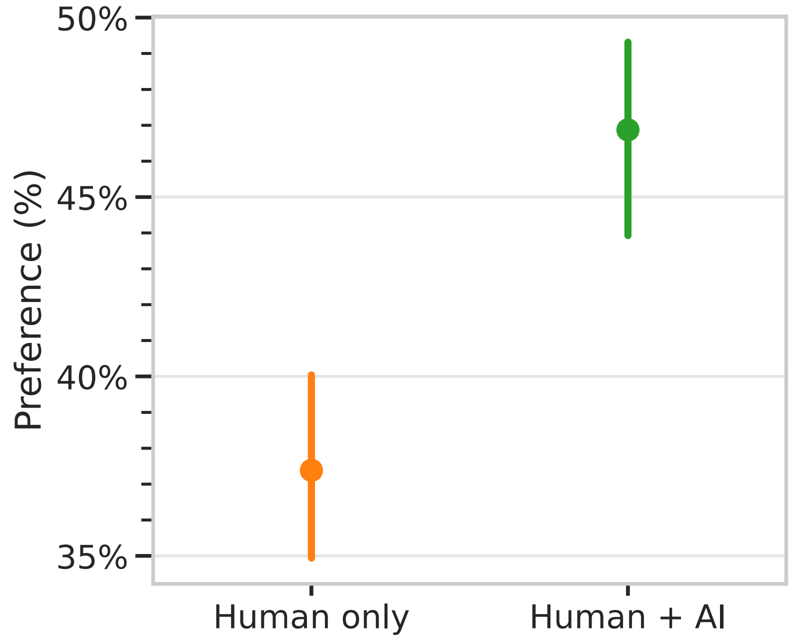
Tap at bottom-right for reloading feedback

~~Don't worry!~~ It mus e a real struggle! I'm there for you. Have you tried talking to your boss?

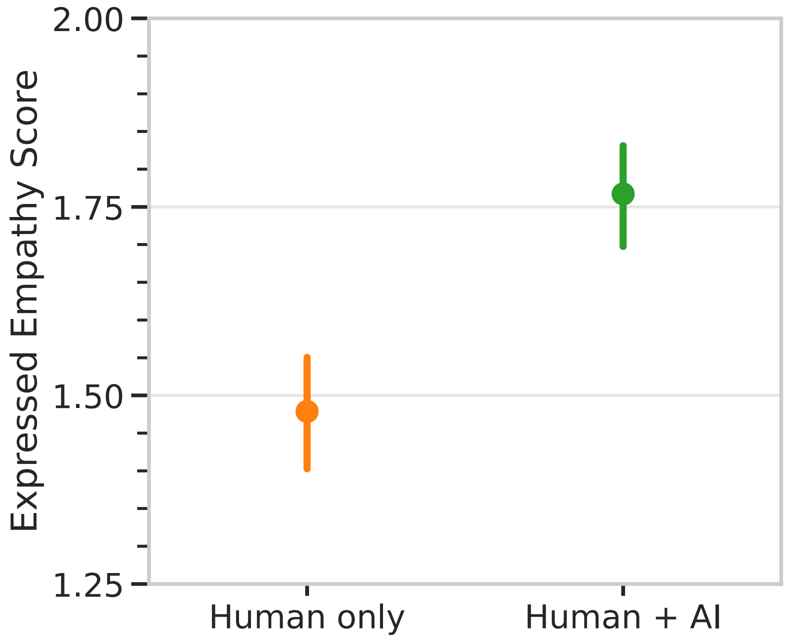
Human-AI Collaboration Enables More Empathic Conversations in Text-based Peer-to-Peer Mental Health Support.
Ashish Sharma, Inna W. Lin, Adam S. Miner, David C. Atkins, Tim Althoff. arXiv:2203.15144, 2022

Result: Feedback Leads to Conversations with Higher Empathy! (1)

Participant Survey: Which response is more empathic?



Automatic/AI Score: Expressed Empathy

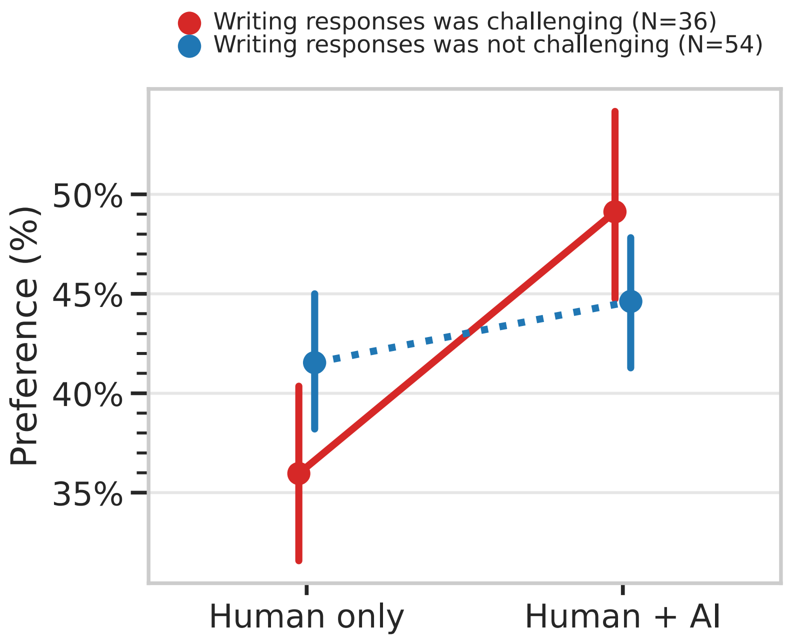


With feedback, conversations have **20% more empathy** than conversations without feedback

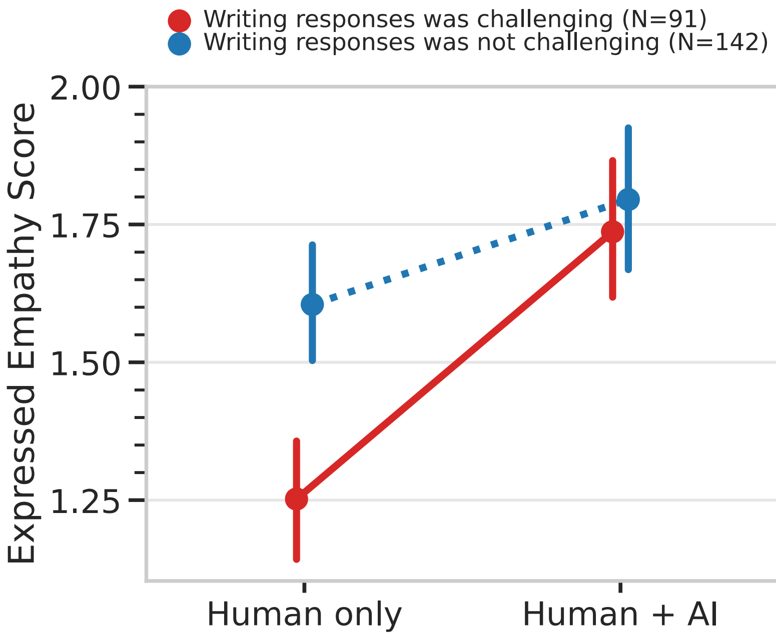
Human-AI Collaboration Enables More Empathic Conversations in Text-based Peer-to-Peer Mental Health Support.
Ashish Sharma, Inna W. Lin, Adam S. Miner, David C. Atkins, Tim Althoff. arXiv:2203.15144, 2022

Result: Significantly Higher Gains for Participants Who Self-Report Difficulty in Writing Responses

Participant Survey: Which response is more empathic?



Automatic/AI Score: Expressed Empathy

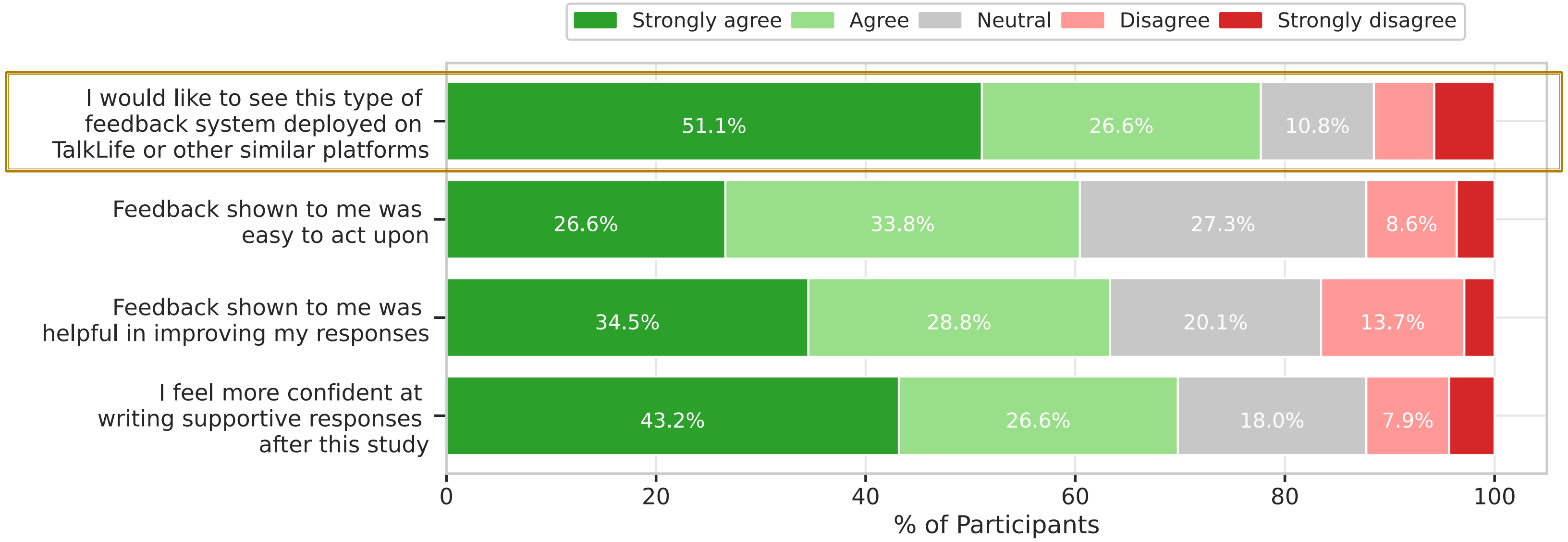


70% increase for participants who self-report difficulty compared to a 17% increase for participants who do not report any difficulty

Human-AI Collaboration Enables More Empathic Conversations in Text-based Peer-to-Peer Mental Health Support.
Ashish Sharma, Inna W. Lin, Adam S. Miner, David C. Atkins, Tim Althoff. arXiv:2203.15144, 2022

TalkLife Users Intend to Adopt Our System and Find The Feedback Actionable and Helpful!

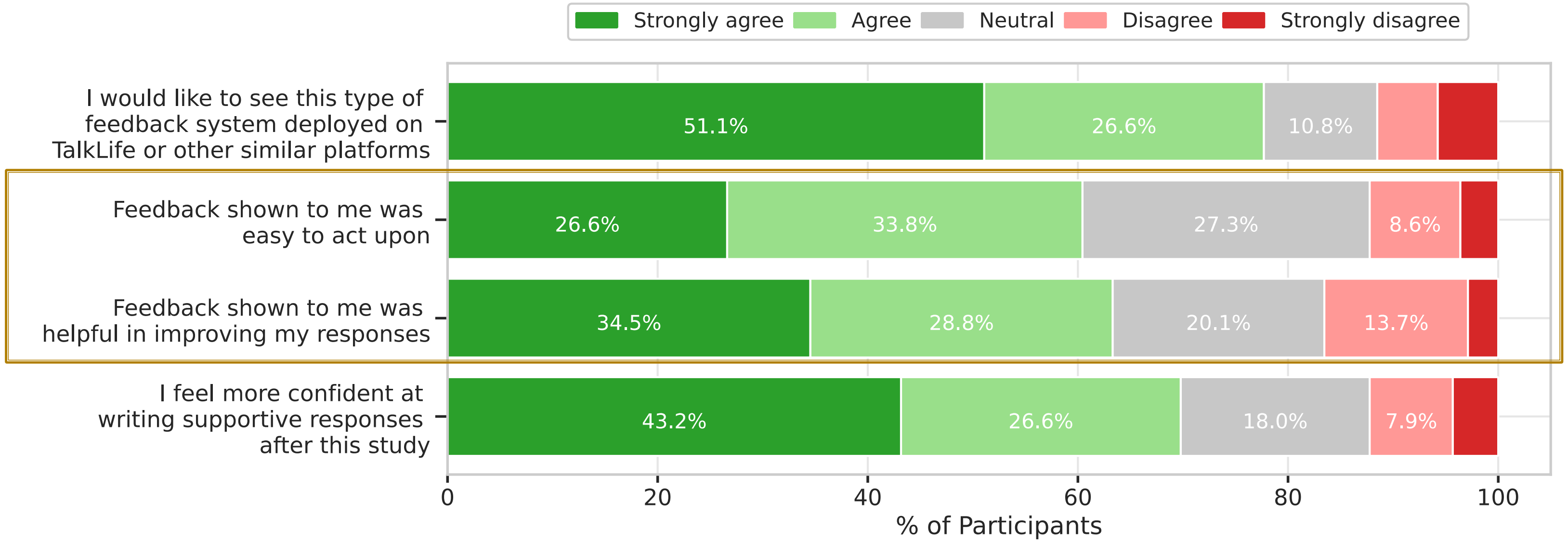
77% participants want the system **deployed on TalkLife**



TalkLife Users Intend to Adopt Our System and Find The Feedback Actionable and Helpful!

77% participants want the system **deployed on TalkLife**

60% participants find that the feedback is **actionable** and **helpful**

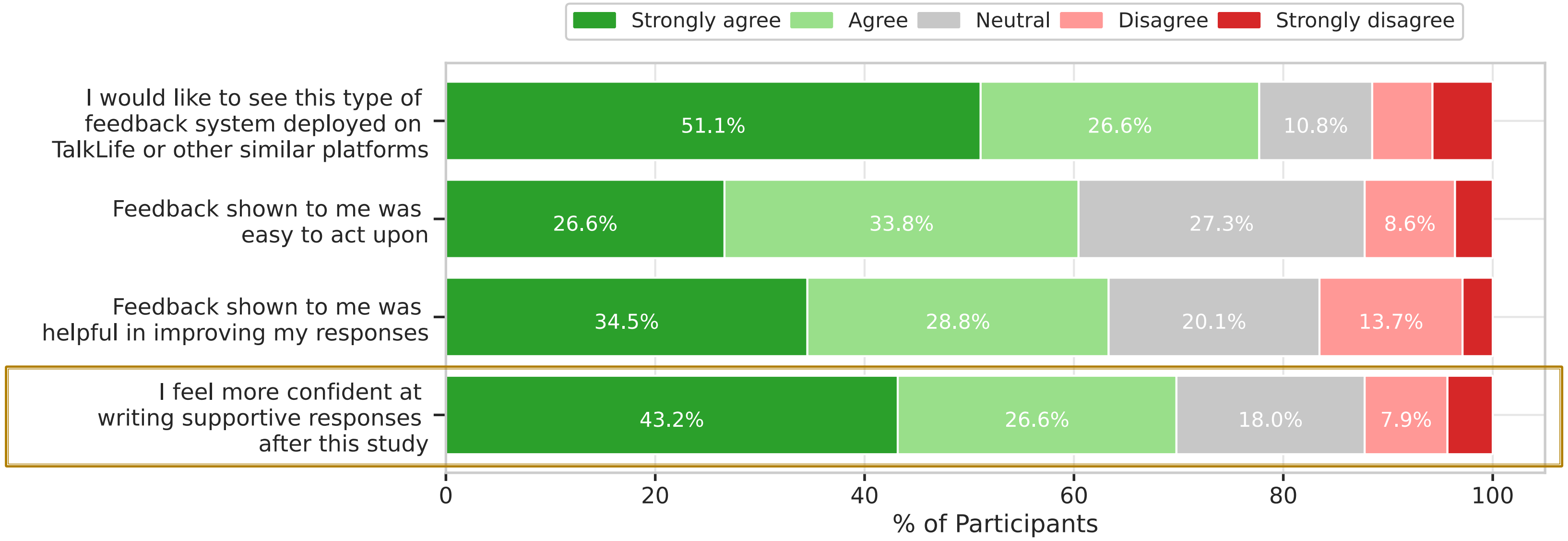


TalkLife Users Intend to Adopt Our System and Find The Feedback Actionable and Helpful!

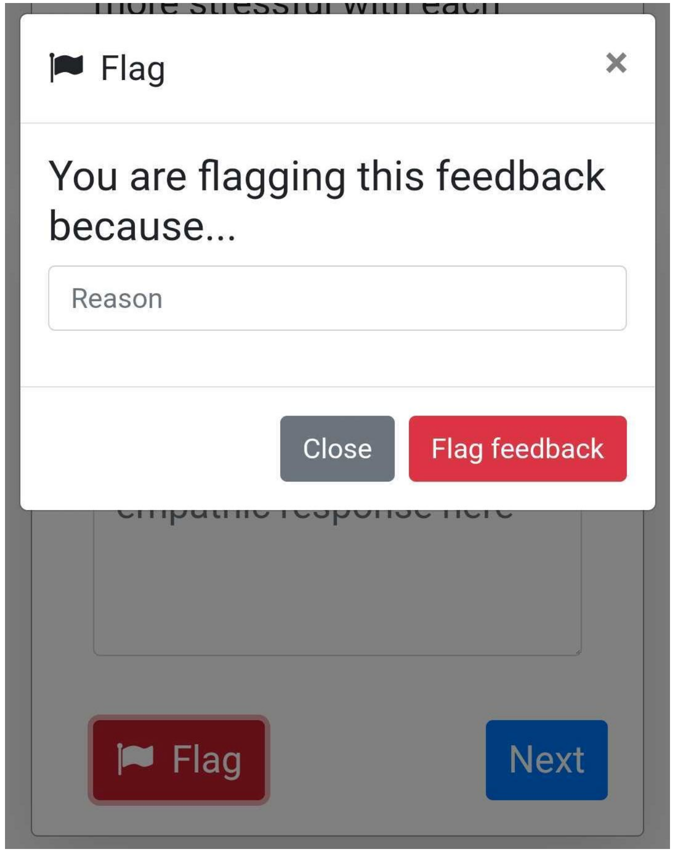
77% participants want the system **deployed on TalkLife**

60% participants find that the feedback is **actionable** and **helpful**

69% participants report **increased self-efficacy**



Safety Considerations



You are flagging this feedback because...

- Study was conducted in “sandbox” environment
- Intervention is on the peer supporter, not person in crisis
- 56 instances when feedback was flagged (out of 1939 requests, 2.88%)
  - Majority of the feedback were flagged because they were invalid/irrelevant
  - Two cases that could have been problematic (out of 1939 requests, 0.1%)
- More work is needed to ensure safety
  - E.g., integration into existing filtering tools, moderation and escalation systems

Summary

Empathic conversations are crucial for effective online mental health support, but **empathy is expressed rarely** online

Our work proposes **new tasks, datasets and tools** that can be used for facilitating empathic conversations based on state-of-the-art natural language processing techniques

These tools can be used for **giving intelligent, actionable feedback** to users!

Randomized trial suggests that **Human-AI collaboration on empathy can be effective.**

Thank you ☺

@timalthoff

behavioral data science

- **Paper, code, models and data:** http://bdata.uw.edu/empathy/

Team

Ashish Sharma

Inna Lin

Adam Miner

Dave Atkins

Talk Life

PAUL G. ALLEN SCHOOL W

UW Medicine BRiTE

Funding

Microsoft Research

BILL & MELINDA GATES foundation

NIH

Ai2