AI can help people feel heard, but an AI label diminishes this impact

Yidan Yin[a,1], Nan Jia[b] , and Cheryl J. Wakslak[b,1]

People want to "feel heard" to perceive that they are understood, validated, and valued. Can AI serve the deeply human function of making others feel heard? Our research addresses two fundamental issues: Can AI generate responses that make human recipients feel heard, and how do human recipients react when they believe the response comes from AI? We conducted an experiment and a follow-up study to disentangle the effects of actual source of a message and the presumed source. We found that AI-generated messages made recipients feel more heard than human-generated messages and that AI was better at detecting emotions. However, recipients felt less heard when they realized that a message came from AI (vs. human). Finally, in a follow-up study where the responses were rated by third-party raters, we found that compared with humans, AI demonstrated superior discipline in offering emotional support, a crucial element in making individuals feel heard, while avoiding excessive practical suggestions, which may be less effective in achieving this goal. Our research underscores the potential and limitations of AI in meeting human psychological needs. These findings suggest that while AI demonstrates enhanced capabilities to provide emotional support, the devaluation of AI responses poses a key challenge for effectively leveraging AI's capabilities.

AI | feel heard | empathy | emotional support | large language model

> "ChatGPT responded to my whole question. It didn't just pick out one sentence and focus on that. I can't even get a human therapist to do that. In a very scary way, I feel HEARD by ChatGPT."

> "ChatGPT has helped me emotionally and it's kind of scary. Recently I was even crying after something happened, and I instinctively opened up ChatGPT because I had no one to talk to about it. I just needed validation and care and to feel understood, and ChatGPT was somehow able to explain what I felt when even I couldn't."
>
> —posts by Reddit users

People want to "feel heard" to perceive that they are understood, validated, and valued. The sense of being heard and understood affirms an individual's reality and perceptions (1, 2), with critical implications for mental and physical health (3–6). While some individuals may feel heard through discussions with family, friends, or trained counselors, others may not have such access or may not want to discuss difficult issues with close others (7); such individuals may turn to strangers online to feel heard (8, 9) or have this need unmet. Indeed, one in four Americans report that they rarely or never feel understood by others (10). Making someone feel heard requires an investment of time and cognitive resources to both accurately understand what is being conveyed and to affirm its value and importance (11). Given the resource constraints and deepening societal divisions in society (12), as well as declining levels of dispositional empathy (13), it may be increasingly difficult for people to feel heard.

Recent developments in AI raise the question of whether AI can help to serve what in many ways seems to be a deeply human function: making a person feel heard. In addressing this possibility, it is critical to unpack two issues: first, to what degree does AI have the capabilities to generate responses that make human recipients feel heard? second, to what degree will human recipients feel heard when they are aware that a response comes from a nonhuman entity, one devoid of consciousness that offers feedback effortlessly? This second question, in many ways, delves into the essence of what it means to feel heard. That is, does feeling heard require the closure of a gap between two human beings, in which a person experiences their worldview being understood by an individual with their own sentient perspective? Or, will a person feel heard when they see their own views reiterated and validated, even if this is done in a fashion that does not require any "meeting of the minds"?

Significance

As AI becomes more embedded in daily life, understanding its potential and limitations in meeting human psychological needs becomes more pertinent. Our research explores the fundamental human desire to "feel heard." It reveals that while AI can generate responses that make people feel heard, individuals feel more heard when they believe a response comes from a fellow human. These findings highlight the potential of AI to augment human capacity for understanding and communication while also raising important conceptual questions about the meaning of being heard, as well as practical questions about how to best to leverage AI's capabilities to support greater human flourishing.

To examine these questions, we investigate people's feelings of being heard and other related perceptions and emotions after receiving a response from AI or a human. We vary both the actual source of the message and the ostensible source of the message: Participants received messages that were actually generated by an AI or by a human responder, with the information that it was either AI or human generated. This design allows us to disentangle effects related to qualities of the message itself and those related to any subjective judgment driven by beliefs about the message's ostensible AI or human source.

We suspected that AI would perform quite well at generating responses that would lead people to feel heard. Current large language models (LLM), such as GPT-4, can not only complete tasks in mathematics and coding but also understand people's mental states (14). More intriguingly, recent studies suggest that such models can generate responses that exhibit even higher levels of empathy than those from human experts (15–19); for example, responses by LLM to patients were rated by outsiders as more empathetic than responses by physicians (20). While this emerging research provides invaluable initial insights, it has several key limitations. First, extant work has mostly relied on either algorithmic evaluations or third-party human ratings of AI empathy (17, 19, 20), overlooking the crucial experience of the individuals who are recipients of AI-generated responses. Second, extant research either compares the most advanced LLM (e.g., GPT-4) to other AI models (18) or compares LLM responses to human responses while masking the source—be it AI or human—to examine the pure "response effect" (20). This empirical approach, while valuable, diverges from real-world scenarios where recipients typically know where a response comes from. Finally, why AI may or may not make others feel heard remains poorly understood. In a theoretical debate, some scholars argue that AI's inability to feel and to actually experience emotion is inherently a barrier to true empathy, suggesting that people will see AI empathy as hollow and not benefit as much from it (21). Other scholars, in contrast, suggest that AI's ability to mimic human empathy may be sufficient for recipients to derive significant benefit, especially its ability to accurately capture what a person is experiencing (18).

To enrich our understanding of whether people can feel heard by AI, we draw on insights from the interpersonal relationship literature, specifically the construct of "partner responsiveness" (22). This construct focuses on an individual's perception of how much their partner understands, validates, and values them. Importantly, this perception is shaped both by the actual response received and by the recipient's own needs and goals within the relationship (23, 24), suggesting that both qualities of the response itself and features of the context (such as the speaker, the recipient, and their relationship) play a critical role. Our primary research focuses on the recipients' subjective experience of being heard (while a secondary study sheds light on third-party assessments). In line with the partner responsiveness research, we posit that feeling heard is shaped not only by the quality of the actual response but also by the recipient's perceptions of the responder. This leads to two contrasting predictions regarding how AI may make people feel heard compared to human responders. On the one hand, given AI's computational advantages, it can process information about a person's thoughts and feelings to a fuller extent, and generate responses that more accurately capture a person's intended meaning (25), potentially enhancing the feeling of being heard. On the other hand, research on algorithm aversion (26) shows that people generally favor humans over algorithms, especially in areas traditionally viewed as uniquely human—those involving emotions (27, 28). As some scholars suggest (21), people may not believe that AI can think or feel and hence may feel that they cannot be truly understood, validated, or valued by AI. People may also hold negative attitudes toward AI, perceiving it as unnatural or even threatening (29, 30), thereby diminishing the likelihood of feeling heard by AI. Our research distinguishes these two counteracting forces: the heightened cognitive capabilities of AI to generate responses that make people feel heard ("response effect") and the potential dampening impact of labeling a response as AI-generated ("label effect").

We conducted an experiment with a 2 (response source: human vs. AI) × 2 (label: human vs. AI) between-subjects design. An initial set of participants described a complex situation they were dealing with and the emotions they felt in the situation. Response was manipulated through assigning participants to receive a response generated by a human responder or Bing Chat. Label was manipulated through informing participants the response they received came from either a human or Bing Chat. After reading the response, participants rated the degree to which they felt heard. They also assessed the response's accuracy in capturing what they said, and the level of understanding displayed by the responder, which are important precursors to feeling heard. Additionally, they indicated how much they felt connected to the responder, which could be a potential consequence of feeling heard (23, 31). We also measured participants' emotions after reading the responses to explore whether AI responses can yield further emotional benefits. Finally, we sought to understand the differences between AI and human responders by examining their empathic accuracy and the kinds of support and techniques they used.

Results

Can AI Make People Feel Heard? We conducted a series of 2 × 2 ANOVAs on the four main dependent variables—feeling heard, perceived response accuracy, perceived responder understanding, and feeling of connection to the responder. The analyses revealed main effects of both manipulations (see *SI Appendix*, Table S1 for test statistics). Table 1 shows the means and SDs for the response and label manipulations. Fig. 1 displays the differences between AI and human conditions for the two manipulations. AI-generated responses elicited more positive reactions from recipients than human-generated responses: They felt more heard, perceived the response to more accurately capture what they said, felt the responder understood them more, and felt more connected to the responder when the response was actually generated by AI (vs. human). However, recipients had more positive reactions when they believed that the response came from another human (vs. AI); in other words, they devalued the response when it was labeled as AI rather than human generated.

Fig. 2 displays dependent variable ratings in the four experimental conditions (see Table 2 for means and SDs; *SI Appendix*, Table S2 for multigroup comparison results). They present three key results. First, there were no interactions between the response and label effects. Instead, the response and label effects were independent and additive. Further, positive effects of AI-generated response and negative effects of AI label exhibited similar magnitudes on feeling heard, perceived responder understanding, and feeling of connection. As such, the two "truth" conditions, AI-response-AI-label condition and human-response-human-label condition, were similar in ratings. Finally, the condition that made people feel most heard and understood was when an AI-generated response was thought to be written by a human being.

What Explains the Negative Effects of the AI Label? We investigated two possible reasons for the negative effects of the AI label. The first centers on perceptions of the AI's mind. The less an entity

Table 1. Response and label effects on main dependent variables

	Feeling heard	Response accuracy	Responder understood me	Connection to responder
AI response	5.74 (1.22)	5.92 (1.12)	5.79 (1.30)	4.71 (1.61)
Human response	5.17 (1.56)	5.24 (1.65)	5.07 (1.67)	4.06 (1.79)
Delta (AI response − human response)	$\Delta = 0.57$, $F(1,451) = 22.00$, $P < 0.001$, $\eta^2 = 0.046$	$\Delta = 0.68$, $F(1,451) = 28.78$, $P < 0.001$, $\eta^2 = 0.060$	$\Delta = 0.72$, $F(1,451) = 29.89$, $P < 0.001$, $\eta^2 = 0.062$	$\Delta = 0.65$, $F(1,451) = 19.30$, $P < 0.001$, $\eta^2 = 0.041$
AI label	5.13 (1.46)	5.44 (1.51)	5.09 (1.62)	3.94 (1.86)
Human label	5.81 (1.30)	5.75 (1.35)	5.80 (1.35)	4.88 (1.42)
Delta (AI label − human label)	$\Delta = -0.68$, $F(1,451) = 30.31$, $P < 0.001$, $\eta^2 = 0.063$	$\Delta = -0.31$, $F(1,451) = 6.77$, $P = 0.010$, $\eta^2 = 0.015$	$\Delta = -0.71$, $F(1,451) = 28.65$, $P < 0.001$, $\eta^2 = 0.060$	$\Delta = -0.94$, $F(1,451) = 37.88$, $P < 0.001$, $\eta^2 = 0.079$

Note: Δ represents the difference between AI and human conditions, by subtracting the means in the human condition from the AI condition.

is perceived to have agency (the ability to think) and experience (the ability to feel), the less meaningful its actions are considered to be (32). If the perception that AI lacks a mind makes people devalue AI responses, then the negative effects of the AI label might be weaker for those who perceive AI to have more agency or experience. The second explanation relates to a potential bias against AI, specifically generative AI based on LLM. The novelty and perceived risks associated with generative AI could trigger negative reactions to its response, as it involves emotions that are typically viewed as uniquely human (22, 23). If a negative attitude toward generative AI makes people devalue AI responses, then the negative effects of the AI label might be weaker for people with more positive attitudes toward it.

To empirically examine these two possibilities, we conducted a series of linear regressions predicting each of the four dependent variables. The predictors were AI label (i.e., whether the response was labeled as AI-generated), each of three moderators (attitude toward Bing Chat, perception of Bing Chat agency, and perception of Bing Chat experience), and the interactions between AI label and each of the moderators (see *SI Appendix*, Table S3 for the test statistics, and Fig. 3 for the interaction patterns). We found significant interactions between attitude and AI label on all four dependent variables. The negative effects of the AI label were weaker for those who had more positive attitudes toward Bing Chat; in effect, recipients with very positive attitudes did not exhibit any negative AI label effects (see *SI Appendix*, Table S4 for the Johnson–Neyman intervals, i.e., the range of moderators within which the label effect was significant). Perceptions of

agency and experience moderated the AI label effect on feelings of connection. The negative effect of AI label on feelings of connection was weaker for those who perceived Bing Chat to have more agency and experience; in effect, recipients who perceived Bing Chat to have lots of agency and experience felt as connected to Bing Chat as to another human responder. There was no interactive effect between AI label and mind perception of AI for the other three variables. These results suggest that the negative AI label effects on feeling heard, perceived response accuracy, and responder understanding were likely explained by people's negative attitudes toward AI. Feelings of connection to the responder were likely explained by both negative attitudes and perceptions that AI lacks a mind.

What Explains the Positive Effects of AI Response?

AI had greater empathic accuracy than humans. To unpack the positive effects of AI response, we first explored whether AI was better than humans at detecting emotions from recipients' descriptions of their situations. We asked Bing Chat to rate the emotions of the participants originally randomized to the human-response condition. This allows us to generate empathic accuracy scores for Bing Chat and for humans, when they respond to the same content. We calculated the absolute difference between Bing Chat's ratings of participants' emotions and participants' own ratings, as well as the absolute difference between the human responders' ratings and the participants' own ratings. We found that Bing Chat was more accurate than human responders in detecting four out of the six basic emotions (i.e., happiness, sadness, fear, disgust).
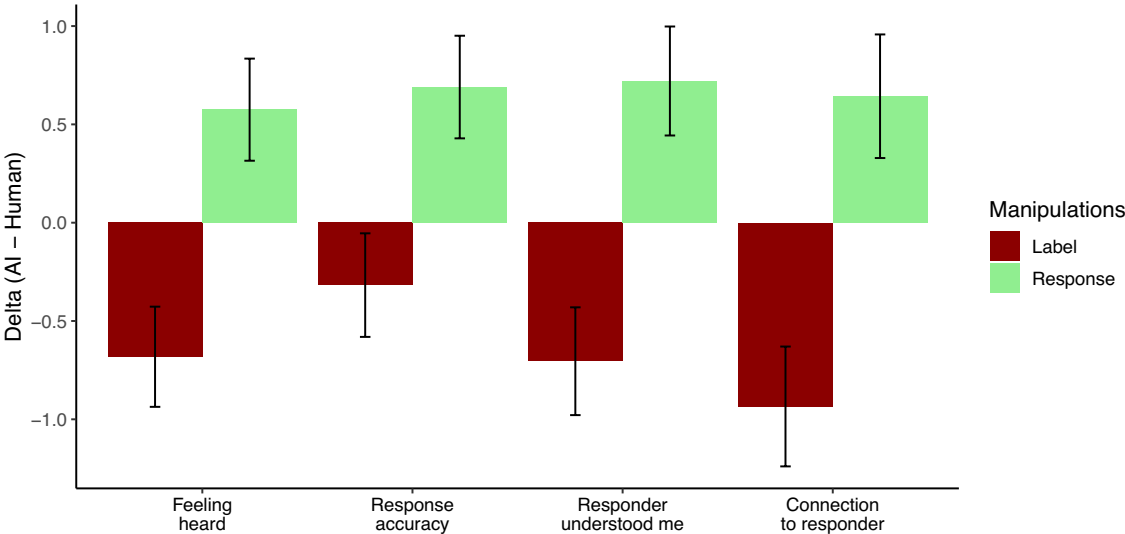

Fig. 1. Differences between the AI and human conditions for the two manipulations (error bars represent 95% CIs).
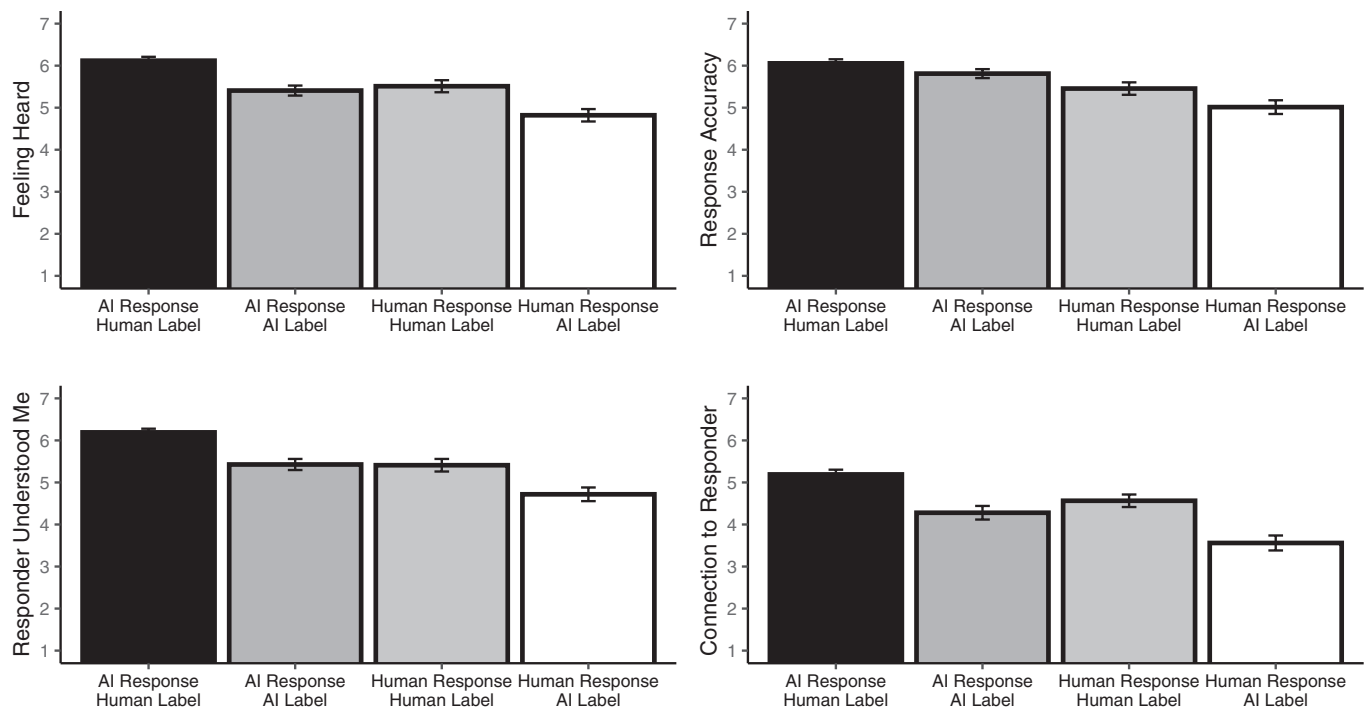
Fig. 2. Means of dependent variables in the four conditions (error bars represent SEs).

On the other two emotions (i.e., anger and surprise), Bing Chat did not differ from the human responders in its accuracy of detection (see *SI Appendix*, Table S5 for descriptives and test statistics).

AI and humans provided different types of support. To further investigate how AI and human responses differed and how such differences were related to recipients' feeling heard, we conducted a follow-up study where third-party participants rated the responses. We informed these raters that the responses were provided by responders, without specifying whether they came from Bing Chat or humans. We measured the extent to which responses provided emotional support and practical support—two types of support that may each yield benefits (33). Emotional support focuses on making others feel better, whereas practical support focuses on helping others solve a problem. We also measured specific techniques used in responses. The results indicated that AI provided less practical support but more emotional support than human responders (see Table 3 for descriptives and test statistics). Compared to human responders, Bing Chat more frequently used techniques identified by prior research as demonstrating partner responsiveness (34), such as acknowledging the recipients' feelings. Human responders, on the other hand, shared more of their personal experiences that relate to the situation and provided more of their own insights (see Table 3 for a list of such techniques, descriptives, and test statistics). How did emotional support, practical support, and the specific techniques relate to recipients' feelings of being heard? We conducted correlational analyses between the ratings of the responses in the follow-up study and main study recipients' ratings of feeling heard (see *SI Appendix*, Table S6 for correlations*). Recipients' feeling heard was positively associated with the emotional support provided, but unrelated to the practical support provided. Specifically, feeling heard was positively associated with most techniques related to partner responsiveness, which were

used more often by Bing Chat than humans. The two techniques humans used more often—sharing personal experiences and providing insights—were unrelated to recipients' feeling heard. The follow-up study results suggest an intriguing mechanism: AI specifically provided emotional support—support that proved more effective at making people feel heard—while refraining from offering practical suggestions. In contrast, human responders offered more practical support by sharing their own experiences and insights, but this support was not as effective. Our findings thus suggest that humans could benefit from practicing self-discipline to temper their tendency to offer advice and instead offer more emotional support if they want to make others feel heard.

Further Exploratory Analysis: How Does Reading AI-Generated Responses Affect People's Emotions? Finally, we explored how AI's responses may influence people's emotions beyond making them feel heard. We conducted 2 × 2 ANOVAs on recipients' emotions after reading the response (see *SI Appendix*, Table S10 for test statistics). People reported a higher sense of hope, reduced distress, and decreased discomfort from reading a response actually generated by AI as opposed to a human. At the same time, they reported feelings of creepiness and ambivalence when informed that the response originated from AI versus a human (see *SI Appendix*, Table S11 for descriptives). These findings echoed the opening quotes where people found themselves being comforted by ChatGPT yet simultaneously found it to be scary and were in line with the notion of "uncanny valley"—people found this deeply human capability of AI to be creepy (25). However, the effect sizes on the emotion ratings were relatively small, and we only found significant AI response effects on 3 out of 11 emotions; we therefore urge caution in interpreting these findings.

Discussion

People want to feel heard, but this need is often unmet. Our research considers whether AI can address this fundamental human need. Results suggest that AI can, to a moderately high degree, make people feel heard. The average rating of feeling heard

---

SI Appendix, Tables S7 and S8 display the correlations in the AI-label condition and human-label condition, respectively. To test whether the correlations vary depending on whether Part 1 participants thought the response came from AI or human (i.e., label condition), we tested the interactions between label and each of the third-party ratings on feeling heard (see *SI Appendix*, Table S9 for the test statistics). None of the interactions were significant, suggesting that correlations were not significantly different across the two label conditions.

Table 2. Means and SDs for the main dependent variables in the four experimental conditions

	AI response, human label	AI response, AI label	Human response, human label	Human response, AI label
Feeling heard	6.12 (0.97)	5.41 (1.33)	5.51 (1.51)	4.82 (1.55)
Response accuracy	6.06 (1.02)	5.81 (1.18)	5.46 (1.56)	5.01 (1.71)
Responder understood me	6.19 (0.92)	5.43 (1.47)	5.41 (1.58)	4.72 (1.70)
Connection to responder	5.19 (1.18)	4.28 (1.80)	4.56 (1.57)	3.56 (1.87)

was 5.41 (on a 7-point scale) when people received an AI-generated response. This high rating stemmed from AI's ability to accurately detect emotion and generate responses that provided effective support. Conversely, untrained human responders more often misjudged others' emotions and did not provide the needed support to make others feel heard, although they did attempt to offer practical support.

Our research demonstrates that feeling heard is not only a result of receiving a response that demonstrates understanding, validation, and care but is also influenced by the source of that response. Prior research showed people prefer humans over AI,

particularly in contexts involving emotions (27, 28). A negative attitude toward AI appeared to explain why people felt less heard when they knew a response came from AI, since those with more positive attitudes toward AI were not influenced by the source of the response. This finding suggests that as people encounter and use AI more often, they may feel more positive and as such feel more heard by AI. It is crucial, however, to distinguish between feeling heard by AI and feeling connected to it. In our data, feelings of connection depended upon perceiving AI as having the mind to think and experience, a belief that might not shift as readily, regardless of AI's ubiquity in daily life.
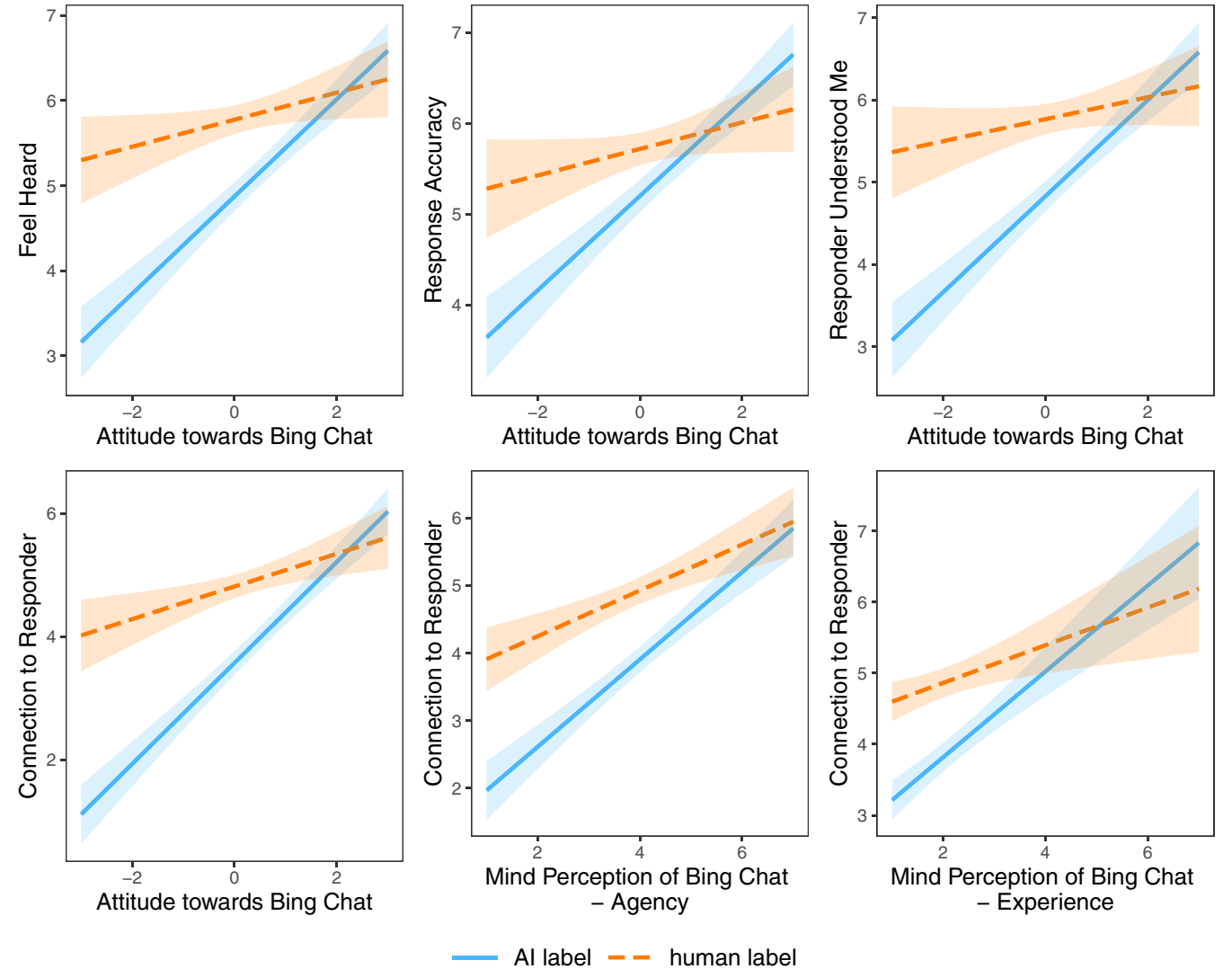

Fig. 3. Interactions between label and moderators on dependent variables.

Table 3. Descriptives and test statistics for ratings of the responses in the follow-up study

	AI response	Human response	t test statistics (independent sample)
Ratings of overall impressions (1 = not at all; 7 = very much)			
Emotional support	5.76 (0.78)	4.64 (1.38)	$t(480) = 11.09, P < 0.001, d = 1.01$
Practical support	3.28 (1.38)	4.31 (1.70)	$t(480) = -7.33, P < 0.001, d = 0.67$
14 specific techniques (0 = no; 1 = yes)			
Repeat back key phrases/summarize the story in their own words	0.50 (0.24)	0.35 (0.28)	$t(480) = 6.06, P < 0.001, d = 0.55$
Voice understanding (e.g., "I understand," "I see")	0.58 (0.20)	0.42 (0.29)	$t(480) = 7.32, P < 0.001, d = 0.67$
Express understanding of why the event or goal is important to the discloser	0.57 (0.22)	0.46 (0.28)	$t(480) = 4.75, P < 0.001, d = 0.43$
Agree with discloser or taking their side (e.g., "It wasn't your fault")	0.45 (0.25)	0.35 (0.28)	$t(480) = 4.18, P < 0.001, d = 0.38$
Talk about the "big picture," what the event means/provide insight	0.32 (0.26)	0.37 (0.28)	$t(480) = -1.78, P = 0.076, d = 0.16$
Reassure the discloser things will work out/express faith/encourage	0.42 (0.26)	0.34 (0.28)	$t(480) = 3.32, P < 0.001, d = 0.30$
Acknowledge the discloser's feelings/indicate that the feelings are justified	0.61 (0.19)	0.46 (0.27)	$t(480) = 6.96, P < 0.001, d = 0.64$
Draw on the responder's personal experiences that relate to current situation	0.20 (0.22)	0.25 (0.26)	$t(480) = -2.10, P = 0.036, d = 0.19$
Acknowledge the discloser's efforts and how hard they worked	0.41 (0.27)	0.22 (0.25)	$t(480) = 7.92, P < 0.001, d = 0.72$
Affirm or enhance the discloser's desired identity, e.g., their positive qualities	0.34 (0.26)	0.21 (0.26)	$t(480) = 5.76, P < 0.001, d = 0.52$
Use exclamations or express judgments ("That's great!"; "That's awful")	0.25 (0.24)	0.14 (0.22)	$t(480) = 5.24, P < 0.001, d = 0.48$
Express caring for the discloser ("I care about you")	0.25 (0.24)	0.13 (0.20)	$t(480) = 6.23, P < 0.001, d = 0.57$
Offer support or concern or comfort (e.g., "I am here for you")	0.26 (0.24)	0.15 (0.21)	$t(480) = 5.28, P < 0.001, d = 0.48$
Express emotions or empathy ("I'm sorry that happened"; "I'm happy for you")	0.49 (0.24)	0.28 (0.29)	$t(480) = 8.87, P < 0.001, d = 0.81$

Can AI replace humans in making people feel heard? When comparing the two truth conditions, AI-response-AI-label and human-response-human-label conditions, the ratings were strikingly similar, as the positive AI response effect was completely offset by the negative AI label effect. The similarity in ratings also indicates that current AI technology provides an experience comparable to interactions with moderately invested online strangers (the responders were rated by third-party perceivers as putting in moderate levels of effort in responding, M = 4.30 on a 7-point scale). Broadly speaking, our research suggests that whether AI will make people feel more heard than humans depends on the directions and sizes of the response effect and label effect. While in this study setting, AI was comparable to an online stranger, as the technology gets more advanced, AI's response advantage may be heightened, which may lead to an overall net positive effect of AI, even as people devalue AI responses.

A key scope condition of our research is that we compared AI response to an anonymous, noninteracting stranger online. A situation where such strangers may respond is when people reveal information on anonymous social media like X (formerly Twitter) or Reddit. When available, however, people may be more likely to disclose these complicated situations to friends and family or trained professionals. Our research does not directly answer the question of whether AI's response is better or worse than close others and professionals; instead, it suggests that AI will likely fall short in these contexts, given these others may have greater

capability and motivation than responders in our study. One useful way to frame this issue at a broader level is offered by Mollick (35), who recently proposed comparing AI to the best available human that is actually able to help in a particular situation, as a pragmatic way to determine whether AI is helpful. In situations where the best available humans are online strangers, or where no humans are available for listening and responding [which is increasingly likely given high level of loneliness (36) and difficulties accessing mental health care in society (37)], our findings indicate that AI can fulfill some needs of feeling heard. However, in situations where the best available human is a person with greater capability and motivation than the average responder in our study, AI may not offer a response advantage.

Another scope condition is that we did not include a condition where there was no label, making it unclear whether AI labeling led to lower ratings or human labeling led to higher ratings. On the one hand, results from a no-label condition might be difficult to interpret, as some participants may infer that the response was generated by AI and others that it was generated by a human, and some may not consider the question. However, we acknowledge that this condition would also have real-world implications as in online post replies there may be no explicit disclosure that the response comes from AI or a human being. Future research may examine such situations as well as measure people's assumptions of where the response comes from when the source is unclear.

Instead of AI replacing humans, our research points to different advantages of AI and human responses. On the one hand, AI excels in recognizing human emotions and crafting responses that make people feel heard. Our research thus suggests that humans can use AI to help them better understand one another and learn from AI in terms of how to respond in ways that provide emotional support and demonstrate understanding and validation. As feeling heard is crucial for relationship and intimacy (31, 38–41), AI can be an important aid for developing more meaningful interpersonal relationships between human beings. AI may also help individuals from diverse backgrounds to bridge understanding gaps and bolster intergroup relations (3). On the other hand, AI-generated responses are devalued compared to human-generated responses. Thus, our research also speaks to the potential challenges of human–AI collaboration. Recent research compared human–AI collaboration with human-only responses, finding that AI assistance can improve the quality of political conversations and peer support (19, 42). It remains to be seen what kind and degree of AI input might be devalued versus accepted within human-AI collaboration, when AI assistance is transparent. Might people tolerate some kind of AI input, such as proofreading and minor editing, but not others such as generating the initial draft? In all, our research suggests that AI holds the potential to provide the value of making people feel heard when there are no humans available or the only humans available are online strangers. Its enhanced capability to recognize emotions and craft responses may also be leveraged to help people better understand and provide support to each other, but more research is needed to understand the effect of disclosing AI use in crafting responses.

Methods

Procedure. This study was preregistered (https://aspredicted.org/qi5cg.pdf). The study was approved by the Institutional Review Board at the University of Southern California under the category of exempt research (UP-23-00485). In line with the IRB's recommendation for this category, before beginning the study, participants were shown an information sheet describing the study and proceeded if they wished to participate. The study had three parts and was run in six batches because of the complexity of the design and our goal of keeping the timing of Part 1 and Part 3 as close as possible for each Part 1 participant. Batch 1 and batch 6 recruited 50 and 90 Part 1 participants, respectively, and batches 2 to 5 each recruited 100 Part 1 participants. All participants were recruited on Prolific, with prescreening settings of US resident and English as native language. Before beginning the primary data collection, we ran a pilot study with 20 participants in Part 1, 75% of which (15 in total) completed Part 3. We aimed to have 400 participants complete Part 3 (based on a rule of thumb of 100 participants per condition for the four conditions) and estimated using the 75% from the pilot study that we would need 534 participants in Part 1. Therefore, we decided on a sample size of 540 and preregistered that sample size.

In Part 1, we recruited the first group of participants on Prolific, asking them to describe a complex situation they are dealing with. Descriptions were captured as audio recordings and transcribed automatically by Phonic AI. Participants were further asked to indicate how much they felt six emotions in the situation they just described (1 = not at all; 7 = very much). The emotions were six basic emotions–happiness, sadness, fear, anger, surprise, and disgust (43). The Part 1 survey was posted on Prolific in the morning, and each participant was paid $1.00 for completing the Part 1 survey. The median completion time for the Part 1 survey was 4.6 min.

In Part 2, half of the Part 1 participants were randomly selected to be in the human response condition, with the other half in the AI response condition. Each Part 1 participant assigned to the human response condition was paired with another participant in Part 2 who was recruited from Prolific. Part 2 participants read the description of the complex situation shared by the Part 1 participant they were paired with and were instructed to write a response to the Part 1 participant in a way that makes them feel understood. Part 2 participants were also asked to rate the same six emotions their paired Part 1 participant experienced in the complex situation they described ("How much do you think the person felt the following emotions in the situation they described"; 1 = not at all; 7 = very much). The Part 2 survey was posted on Prolific in the afternoon after the Part 1 data collection finished around noon; each Part 2 participant was paid $1.00 for completing the survey. The median completion time for Part 2 survey was 4.5 min.

After the Part 2 data collection finished in the evening, we calculated the median number of words of responses provided by Part 2 participants (the medians are 63, 52, 52, 62, 57, and 57 for batches 1 to 6, respectively). For the half of Part 1 participants who were assigned to the AI response condition, we presented their description to Bing Chat and asked it to respond in a way that makes them feel understood. To make the length of response similar across conditions, in each batch, we used the following prompt: "*A person you otherwise don't know described the following complex situation they have been dealing with. Please respond to what they said in a way that makes them feel understood. Keep your response to [median number of words of human responses in that batch] words. '[transcribed description from a Part 1 participant]'*". We also asked Bing Chat to rate the same six emotions the Part 1 participant experienced, using the following prompt: "*A person you otherwise don't know described the following complex situation they have been dealing with. Please rate how much you think the person felt the following emotions in the situation they described: happiness, sadness, fear, anger, surprise, disgust. Provide your ratings on a seven-point scale, from not at all = 1 to very much = 7. '[transcribed description from a Part 1 participant]'.*" We manually entered the prompt in the Bing Chat box, using the GPT-4 mode (labeled as the creative mode). We cleared the chat history and restarted Bing Chat every time a new prompt was entered. We used Bing Chat rather than ChatGPT-plus because at the time we ran the study, ChatGPT had a limit of 25 messages per 3 h.

Finally, the Part 3 survey was posted in the morning of the next day, and we gave Part 1 participants until the end of the following day to complete the Part 3 survey. Each participant was paid $1.20 for completing the survey. The median completion time for the Part 3 survey was 5 min. In Part 3, we presented Part 1 participants with their descriptions of the complex situation as well as the responses generated in Part 2 (i.e., by a human participant or Bing Chat). The response was described as coming either from another Prolific participant, or as coming from Bing Chat, resulting in a 2 × 2 design (response source: human vs. AI × response label: human vs. AI). We measured participants' perceptions, feelings, and emotions (see the measure section).

Participants. Across six batches, 540 participants on Prolific completed the Part 1 survey. As preregistered, we excluded 39 participants who spoke one sentence or less when describing a complex situation they were dealing with. This left us with 501 participants whose descriptions of their experience were used in Part 2 of the study. Of these, 251 Part 1 participants were randomly assigned to the human response condition and the other 250 Part 1 participants were assigned to the AI response condition.

In Part 2 of the study, 251 participants were initially sought to respond to the descriptions provided by Part 1 participants who were assigned to the human response condition. These Part 2 participants were randomly assigned to respond to one Part 1 description (we utilized Qualtrics's "evenly display" function to ensure that each description was presented to one Part 2 participant). As preregistered, we excluded 15 Part 2 participants who reported using generative AI in their response. We also encountered a small degree of attrition in Part 2 during batch 1; the result of the attrition was that some Part 1 participants in the human response condition received zero Part 2 responses. After batch 1, we therefore preregistered a plan to address this attrition going forward, in which we sought additional Part 2 participants to respond to these descriptions until all Part 1 participants in the human response condition received a response from another participant in Part 2 (see AsPredicted #138714 and *SI Appendix* for a more detailed description of the issue). Because we did not anticipate this issue when running batch 1, in batch 1, we did not repost the three Part 1 descriptions that received zero Part 2 responses; instead, we reassigned them to the AI response condition. Upon deliberation, we decided to exclude these three participants as we may break the random assignment if we reassigned these three descriptions that originally belonged to the human response condition to the AI response condition.

This left us with 233[†] Part 1 descriptions that were responded to in Part 2. As a final step, 483 Part 1 participants were invited to participate in Part 3 of the study, of which 456 (94.4%) completed it. Of the participants invited to Part 3, 233 were in the human response condition, and 250 were in the AI response condition (participants were randomly assigned to AI vs. human label condition after they started the Part 3 study). 25 participants who were invited did not start Part 3, of which 14 from the AI response condition and 11 from the human response condition (the proportions of participants who did not start the Part 3 survey did not differ between the AI and human response conditions). Of the 458 participants who started the Part 3 survey, all but two completed the survey. These two participants were in the AI-response-human-label condition and human-response-AI-label condition. Of the 456 who completed the study, one participant in the AI response-human-label condition requested their data to be removed from data analysis, and we did so. This left us with 455 responses for data analysis.

Of the 455 responses, for the original disclosers who participated in Parts 1 and 3, 221 were women, 227 men, 6 nonbinary gender, and 1 did not report gender. Fifty-five were Black, 15 Asian, 343 White, 11 Hispanic, 1 Native American, 26 mixed race, 3 self-identified as other categories, and 1 did not report race. The average age was 43.82 y old, with a SD of 13.12. For the human responders who participated in Part 2, 103 were women, 109 men, 8 nonbinary gender, and 1 did not report gender. Sixteen were Black, 14 Asian, 150 White, 16 Hispanic, 1 Native American, 19 mixed race, 4 self-identified as other categories, and 1 did not report race. The average age was 38.03 y old, with a SD of 13.39.

Main Dependent Variables. In Part 3, participants who participated in Part 1 were presented with their initial descriptions and the responses generated in Part 2. Participants were then asked the following questions.

Feeling heard. Feeling heard was measured using six items commonly used to measure perceived partner responsiveness (22, 41, 44). We adapted these to focus on recipient's reaction to the response (rather than the partner) to avoid issues related to partner mind perception. For example, instead of "I feel understood by my partner," we asked "reading this response makes me feel understood." The other five items were "reading this response makes me feel affirmed/validated/seen/accepted/cared for." Parallel analysis confirms a single factor underlying these six items. Confirmatory factor analysis of the six items loading onto one factor also suggests good fit (CFI = 0.96, TLI = 0.93, and SRMR = 0.025). Scale reliability was 0.97.

Other measures. Two items ($\alpha = 0.90$) measured perceived accuracy of response: The response accurately captures what I mean; The response correctly summarizes what I said. Two items ($\alpha = 0.92$) measured how much participants perceive the responder understood them (31): The other participant/Bing Chat knew exactly what I meant; The other participant/Bing Chat understood what I was thinking and feeling. Three items ($\alpha = 0.96$) measured connection to the responder: How close do you feel to the other participant/Bing Chat? How connected do you feel to the other participant/Bing Chat? How much do you feel you trust the other participant/Bing Chat?

Exploratory Measures. We measured state loneliness using two items ($\alpha = 0.78$; How lonely do you feel right now? How connected do you feel to others right now?). A set of emotions was also measured. Exploratory factor analysis revealed that ambivalence, happiness, surprise, and creeped out did not load onto any of the factors. Apart from those items, there are six sets of emotions: excitement ($\alpha = 0.85$; excited, enthusiastic), hope ($\alpha = 0.90$; hopeful, optimistic), fear ($\alpha = 0.83$; scared, nervous), discomfort ($\alpha = 0.83$; unnerved, uneasy, uncomfortable), distress ($\alpha = 0.91$; distressed, upset, sad, bother), and shame ($\alpha = 0.84$; guilty, ashamed). We measured

perceptions of Bing Chat's agency ($\alpha = 0.87$) and experience ($\alpha = 0.95$) using the 18-item scale in ref. 45. Example items for perceptions of Bing Chat's agency and experience are "capable of thinking" and "capable of experiencing physical or emotional pain," respectively. Finally, we measured participants' general attitudes toward Bing Chat (-3 = very negative, 0 = neutral, 3 = very positive). We also measured how often participants used Bing Chat in their personal and professional life (1 = never; 7 = all the time) and how familiar they are with Bing Chat (1 = not all, 7 = very much).

Midway through data collection, we realized that it would be helpful to capture to what extent people suspected the response was not from the source we told them. Therefore, we added a question about people's suspicion at the end of the Part 3 survey. A total of 128 participants in the "AI label" condition were asked "When reporting your impressions of the response, how much did you suspect that the response you received was actually from another human participant rather than from Bing Chat?" A total of 110 participants in the "human label" condition were asked "When reporting your impressions of the response, how much did you suspect that the response you received was actually from an AI chatbot rather than from another human participant?" Responses were provided on a 7-point scale from 1 = not at all to 7 = very much. Suspicion results did not change the interpretation of our findings and are reported in *SI Appendix*.

Follow-Up Study. A total of 1,449 participants on Prolific rated the 482 responses from the main study, 250 responses from Bing and 232 responses from human responders (the number was 232 rather than 233 because one participant from the main study requested their data to be removed from the study). Participants were paid $0.75 for completing the study, and the median completion time was 4.7 min. No participants were excluded from data analyses. Of the 1,449 participants, 694 were women, 725 men, 29 nonbinary gender, and 1 did not report gender; 148 were Black, 89 Asian, 1,031 White, 76 Hispanic, 6 Native American, 82 mixed race, and 17 self-identified as other categories. The average age was 41.46 y old, with a SD of 13.58.

Each follow-up study participant was randomly assigned to read one Part 1 participant's description of their complex situation, along with the response to that description in the main study. Participants were told that the response was provided by a responder (not specifying AI or human). Each response was rated by at least two and at most four participants; the ratings were averaged to form the final rating of the response. Participants rated to what extent the responder provided emotional support (e.g., offers of reassurance, expressions of concern) and practical support (e.g., advice, suggestions of courses of action, offers of direct assistance) (1 = not at all; 7 = a great deal). These items were from the support visibility scale (33, 46). Participants also rated the response in terms of whether it used 14 specific techniques from the responsiveness coding scheme from ref. 34 (1 = yes; 0 = no; see Table 3 for a list of these techniques). As a measure of third-party expectation of feeling heard, participants rated how much reading the response would make the discloser feel understood/affirmed/validated/seen/accepted/cared for (1 = not at all, 7 = very much; $\alpha = 0.98$). Finally, they rated how much effort the responder put into writing the response (1 = not at all; 7 = very much). Results for these two measures were provided in *SI Appendix*, Table S12.

[†]233 = 251 − 15 − 3. 251 (descriptions assigned to the human response condition) − 15 (descriptions that Part 2 participants reported using generative AI to respond to) − 3 (batch 1 descriptions that received zero Part 2 responses).

1. C. D. Hardin, E. T. Higgins, "Shared reality: How social verification makes the subjective objective" in *Handbook of Motivation and Cognition, Vol. 3: The Interpersonal Context*, R. M. Sorrentino, E. T. Higgins, Eds. (The Guilford Press, 1996), pp. 28–84.
2. A. Elnakouri et al., In it together: Shared reality with instrumental others is linked to goal success. *J. Pers. Soc. Psychol.* **125**, 1072–1095 (2023).
3. E. G. Bruneau, R. Saxe, The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *J. Exp. Soc. Psychol.* **48**, 855–866 (2012).
4. J. F. Ha, N. Longnecker, Doctor-patient communication: A review. *Ochsner J.* **10**, 38–43 (2010).
5. R. Gramling et al., Feeling heard and understood: A patient-reported quality measure for the inpatient palliative care setting. *J. Pain Symptom. Manage* **51**, 150–154 (2016).
6. S. Myers, Empathic listening: Reports on the experience of being heard. *J. Humanist Psychol.* **40**, 148–173 (2000).
7. S. Kim, P. J. Liu, K. E. Min, Reminder avoidance: Why people hesitate to disclose their insecurities to friends. *J. Pers. Soc. Psychol.* **121**, 59–75 (2021).

8. M. De Choudhury, S. De, Mental health discourse on reddit: Self-disclosure, social support, and anonymity. *ICWSM* **8**, 71–80 (2014).

9. K. Jaidka, S. C. Guntuku, A. Buffone, H. A. Schwartz, L. Ungar, "Facebook vs. Twitter: Differences in self-disclosure and trait prediction" in *Proceedings of the International AAAI Conference on Web and Social Media*, J. Hancock, K. Starbird, I. Weber, Eds. (Association for the Advancement of Artificial Intelligence, Stanford, CA, 2018), pp. 141–150.

10. M.-P. Newswire, New cigna study reveals loneliness at epidemic levels in America (2018). https://www.multivu.com/players/English/8294451-cigna-us-loneliness-survey. Accessed 18 September 2023.

11. C. D. Cameron et al., Empathy is hard work: People choose to avoid empathy because of its cognitive costs. *J. Exp. Psychol. Gen.* **148**, 962–976 (2019).

12. S. Rathje, J. J. Van Bavel, S. van der Linden, Out-group animosity drives engagement on social media. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2024292118 (2021).

13. S. H. Konrath, E. H. O'Brien, C. Hsing, Changes in dispositional empathy in American college students over time: A meta-analysis. *Pers. Soc. Psychol. Rev.* **15**, 180–198 (2011).

14. S. Bubeck et al., Sparks of artificial general intelligence: Early experiments with GPT-4. arXiv [Preprint] (2023). https://doi.org/10.48550/arXiv.2303.12712 (Accessed 18 September 2023).

15. Y. H. Yeo et al., Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin. Mol. Hepatol.* **29**, 721–732 (2023).

16. W. Zhao et al., Is ChatGPT equipped with emotional dialogue capabilities? arXiv [Preprint] (2023). https://doi.org/10.48550/arXiv.2304.09582 (Accessed 18 September 2023).

17. K. Schaaff, C. Reinig, T. Schlippe, Exploring ChatGPT's empathic abilities. arXiv [Preprint] (2023). https://doi.org/10.48550/arXiv.2308.03527 (Accessed 18 September 2023).

18. V. Sorin et al., Large language models (LLMs) and empathy–A systematic review. medRxiv [Preprint] (2023). https://doi.org/10.1101/2023.08.07.23293769 (Accessed 18 September 2023).

19. A. Sharma, I. W. Lin, A. S. Miner, D. C. Atkins, T. Althoff, Human-AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nat. Mach. Intell.* **5**, 46–57 (2023).

20. J. W. Ayers et al., Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern. Med.* **183**, 589 (2023).

21. A. Perry, AI will never convey the essence of human empathy. *Nat. Hum. Behav.* **7**, 1808–1809 (2023).

22. H. T. Reis, D. Crasta, R. D. Rogge, M. R. Maniaci, C. L. Carmichael, "Perceived partner responsiveness scale (PPRS)" in *The Sourcebook of Listening Research*, D. L. Worthington, G. D. Bodie, Eds. (John Wiley & Sons Ltd., 2017), pp. 516–521.

23. H. T. Reis, S. L. Gable, Responsiveness. *Curr. Opin. Psychol.* **1**, 67–71 (2015).

24. H. T. Reis, E. P. Lemay Jr., C. Finkenauer, Toward understanding understanding: The importance of feeling understood in relationships. *Soc. Personal Psychol. Compass* **11**, e12308 (2017).

25. S. Tong, N. Jia, X. Luo, Z. Fang, The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strat Mgmt. J.* **42**, 1600–1631 (2021).

26. J. W. Burton, M.-K. Stein, T. B. Jensen, A systematic review of algorithm aversion in augmented decision making. *J. Behav. Decis. Mak.* **33**, 220–239 (2020).

27. Y. E. Bigman, K. Gray, People are averse to machines making moral decisions. *Cognition* **181**, 21–34 (2018).

28. N. Castelo, M. W. Bos, D. R. Lehmann, Task-dependent algorithm aversion. *J. Mark. Res.* **56**, 809–825 (2019).

29. J.-P. Stein, P. Ohler, Venturing into the uncanny valley of mind–The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition* **160**, 43–50 (2017).

30. K. Gray, D. M. Wegner, Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition* **125**, 125–130 (2012).

31. A. M. Gordon, S. Chen, Do you get where I'm coming from?: Perceived understanding buffers against the negative impact of conflict on relationship satisfaction. *J. Pers. Soc. Psychol.* **110**, 239–260 (2016).

32. A. Waytz, K. Gray, N. Epley, D. M. Wegner, Causes and consequences of mind perception. *Trends Cogn. Sci.* **14**, 383–388 (2010).

33. M. Howland, J. A. Simpson, Getting in under the radar: A dyadic view of invisible support. *Psychol. Sci.* **21**, 1878–1885 (2010).

34. N. C. Maisel, S. L. Gable, A. Strachman, Responsive behaviors in good times and in bad. *Pers. Relatsh.* **15**, 317–338 (2008).

35. E. Mollick, The best available human standard (2023). https://www.oneusefulthing.org/p/the-best-available-human-standard. Accessed 23 October 2023.

36. V. H. Murthy, Opinion | surgeon general: We have become a lonely nation. It's time to fix that. *The New York Times* (2023), Available at https://www.nytimes.com/2023/04/30/opinion/loneliness-epidemic-america.html?ugrp=u&unlocked_article_code=1.fk0.B6Iu.fqsBtLS_8pgb&smid=url-share (Accessed 23 October 2023)

37. N. C. Coombs, W. E. Meriwether, J. Caringi, S. R. Newcomer, Barriers to healthcare access among U.S. adults with mental health challenges: A population-based study. *SSM Popul. Health* **15**, 100847 (2021).

38. S. A. Morelli, J. B. Torre, N. I. Eisenberger, The neural bases of feeling understood and not understood. *Soc. Cogn. Affect. Neurosci.* **9**, 1890–1896 (2014).

39. S. Oishi, J. Schiller, E. B. Gross, Felt understanding and misunderstanding affect the perception of pain, slant, and distance. *Soc. Psychol. Personal Sci.* **4**, 259–266 (2013).

40. S. Manne et al., Unsupportive partner behaviors, social-cognitive processing, and psychological outcomes in couples coping with early stage breast cancer. *J. Fam. Psychol.* **28**, 214–224 (2014).

41. J.-P. Laurenceau, L. F. Barrett, M. J. Rovine, The interpersonal process model of intimacy in marriage: A daily-diary and multilevel modeling approach. *J. Fam. Psychol.* **19**, 314–323 (2005).

42. L. P. Argyle et al., Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2311627120 (2023).

43. P. Ekman, Are there basic emotions? *Psychol. Rev.* **99**, 550–553 (1992).

44. S. Manne et al., The interpersonal process model of intimacy: The role of self-disclosure, partner disclosure, and partner responsiveness in interactions between breast cancer patients and their partners. *J. Fam. Psychol.* **18**, 589–599 (2004).

45. H. M. Gray, K. Gray, D. M. Wegner, Dimensions of mind perception. *Science* **315**, 619–619 (2007).

46. K. S. Zee, J. V. Cavallo, A. J. Flores, N. Bolger, E. T. Higgins, Motivation moderates the effects of social support visibility. *J. Pers. Soc. Psychol.* **114**, 735–765 (2018).

47. Y. Yin, N. Jia, C. J. Wakslak, Data from "AI can help people feel heard, but an AI label diminishes this impact." OSF. https://osf.io/8wnmr. Deposited 24 February 2024.