Journal of Personality and Social Psychology

Judged by Humans, Comforted by Machines: How Psychological Expectations Shape Experiences and Perceptions of AI Support

Roshni Raveendhran and Trevor A. Foulk

Online First Publication, September 3, 2026. https://dx.doi.org/10.1037/pspa0000494

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Personality and Social Psychology:
Attitudes and Social Cognition

ISSN: 0022-3514

https://doi.org/10.1037/pspa0000494

Judged by Humans, Comforted by Machines: How Psychological Expectations Shape Experiences and Perceptions of AI Support

Roshni Raveendhran[1] and Trevor A. Foulk[2]
[1] Darden School of Business, University of Virginia
[2] Warrington College of Business, University of Florida

People frequently encounter minor aversive experiences—such as rude interactions or social slights—that, while not major events, can still affect psychological well-being. In this research, we examine how people experience and perceive social support from artificial intelligence (AI) following such experiences. Drawing on mind perception theory, we propose that initial expectations about AI's experiential capacity relative to humans shape how people experience and perceive AI support. A pilot study demonstrates that people have different initial expectations of AI's versus humans' experiential capacity. Across six studies (five preregistered; total $N = 3{,}341$), we found that these initial expectations produce two distinct effects: People report lower fear of negative evaluation after receiving support from AI versus a human, which in turn improves well-being outcomes (Studies 1–3), and people perceive AI as exceeding their expectations for empathy, which in turn drives perceived support effectiveness (Studies 4–6). These effects emerged across controlled simulations and real interactions with actual human and AI support givers. Together, these findings reveal that initial psychological expectations about AI's experiential capacity play a central role in shaping the experience and perception of AI support.

Statement of Limitations

In this research, we found that people's psychological expectations about AI shape how they experience and perceive AI support in response to minor aversive experiences. While our findings consistently support the theorized model, several limitations warrant attention. First, all studies used employed U.S. adults recruited online, limiting generalizability to other cultural contexts wherein norms around judgment and support seeking may differ. Second, our studies focused on episodic emotional support following minor everyday aversive experiences delivered through digitally mediated channels; we cannot speak to whether these findings extend to more serious emotional challenges, sustained support contexts, or face-to-face interactions. Third, outcomes were measured immediately after the support interaction, and we do not know how long these effects last. Fourth, our theoretical framework assumes that people perceive AI as having lower experiential capacity than humans, but our studies do not disentangle whether benefits stem from perceived capacity limitations or perceived motivational tendencies—a distinction that may become more consequential as AI advances. Fifth, while we identify exceeded empathy expectations as a key mechanism, other factors may also contribute to perceptions of AI support effectiveness. Sixth, all outcome measures were self-report and do not capture behavioral changes in support seeking or coping. Finally, our findings rest on people's current initial expectations about AI, which may shift as people gain more experience with AI support tools, potentially changing the effects we observe.

Keywords: psychological expectations, social support, artificial intelligence, mind perception

Supplemental materials: https://doi.org/10.1037/pspa0000494.supp

Paul Conway served as action editor.

Roshni Raveendhran https://orcid.org/0000-0002-0288-057X

The authors are deeply grateful to Jim Detert for his steadfast encouragement at critical stages of this project. The authors owe a debt of gratitude to Sathish Gangichetty for architecting and developing custom AI chatbots for this project. The authors thank Peter Belmi for his support and insightful conversations and Tami Kim, Gabrielle Adams, members of the PEER Lab at University of Virginia Darden, and University of Virginia Psychology seminar participants for their substantive comments on earlier versions of the article. Finally, the authors thank Corrine Bresky for her invaluable research assistance.

Roshni Raveendhran played a lead role in data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, validation, visualization, and writing–original draft and an equal role in conceptualization and writing–review and editing. Trevor A. Foulk played a supporting role in formal analysis and methodology and an equal role in conceptualization and writing–review and editing.

Correspondence concerning this article should be addressed to Roshni Raveendhran, Darden School of Business, University of Virginia, 100 Darden Boulevard, Charlottesville, VA 22903, United States. Email: raveendhranr@darden.virginia.edu

When people face difficult social experiences such as a rude comment, a painful slight, or a moment of rejection, they are increasingly turning to artificial intelligence (AI) for social support (De Freitas et al., 2025; Fitzpatrick et al., 2017). As AI takes on these support roles, a natural question arises: Can AI provide social support as effectively as humans? This question is particularly intriguing because people perceive AI as lacking the capacity to truly feel, care, or be emotionally invested in a relationship (H. M. Gray et al., 2007; Waytz et al., 2010), the very qualities that make social support effective (Dakof & Taylor, 1990). Existing research offers conflicting answers: Some studies suggest that social behaviors from AI feel unfulfilling or inappropriate (Castelo et al., 2019; Dang & Liu, 2024), while others find that AI interactions in social contexts elicit meaningful responses (De Freitas et al., 2025; Yin et al., 2024). In this research, we draw on mind perception theory to argue that people's initial expectations about AI—specifically, its lower perceived ability to experience emotions (Bigman et al., 2023; H. M. Gray et al., 2007; K. Gray & Wegner, 2012)—create conditions that reshape the support experience after minor aversive experiences, paradoxically improving both well-being and perceived support effectiveness.

Understanding how people experience and perceive AI support matters because the need for accessible social support is acute. People regularly encounter minor aversive social experiences like rude interactions and social slights that, while not traumatic, erode mood and deteriorate well-being (Porath & Pearson, 2013; Schilpzand et al., 2016; Woolum et al., 2024). Indeed, research on daily hassles suggests that the accumulation of such minor stressors is a stronger predictor of psychological distress than major life events (Kanner et al., 1981), underscoring that these experiences, while individually minor, are consequential for well-being. Social support is one of the most effective buffers against the consequences of these experiences (Cutrona, 1990), yet access is limited: Nearly half of American adults report feeling lonely, the share with three or fewer close friends has nearly doubled since 1990 (U.S. Department of Health and Human Services, 2023), and globally, one in five employees regularly feels lonely at work (Gallup, 2024, 2025). Compounding the problem, people often hesitate to seek or offer support—even when others are available—because seekers underestimate others' willingness, helpfulness, or responsiveness, while givers underestimate how their support will be perceived, a well-documented phenomenon known as undersociality (Dungan et al., 2022; Kumar & Epley, 2023). AI systems that can hold interactive conversations (Adiwardana et al., 2020; Breazeal, 2003; Rizzo et al., 2016; Schuetzler et al., 2020) are increasingly filling this gap, providing support that is accessible, immediate, and available without the social barriers that constrain human support. However, people bring beliefs about what AI can and cannot do into these interactions, and we draw on mind perception theory to argue that these initial expectations shape how people experience and perceive AI support.

Mind perception theory suggests that people attribute less mind to AI than to humans, especially a lower capacity to sense, feel, and have emotional and social motivations (Bigman et al., 2023; Bigman & Gray, 2018; Epley & Waytz, 2010; H. M. Gray et al., 2007). We argue that these initial beliefs—the expectations people bring to their early encounters with AI support—produce two distinct psychological effects. First, because people expect AI to lack the social motivation to judge them, they experience lower fear of negative evaluation (FNE) when receiving AI (vs. human) support, leading to improved well-being outcomes after minor aversive experiences. Second, because people set lower baseline expectations for AI's empathy, AI is more likely to exceed those initial expectations when providing support, leading people to perceive AI support as more effective. As we show across a pilot study and six experiments, the same initial expectations about AI's low capacity for emotional experience (compared to humans) that might seem like a liability in other interpersonal contexts can become an asset in social support contexts, inverting the typical logic of undersociality, where low expectations of others produce negative outcomes (Epley et al., 2022).

AI and Social Support

Social support takes many forms, from brief moments of reassurance to long-term therapy, and serves various functions: emotional support (helping someone feel better), instrumental support (providing tangible help or resources), and appraisal support (offering evaluative feedback; Cutrona, 1990). Support can also be episodic (short interactions) or chronic (unfolding over time, such as in therapist–client relationships; Shockley & Allen, 2013). Given this breadth, scholars emphasize the importance of specifying the type of support being examined (Jacobson, 1986; Shockley & Allen, 2013; Veiel, 1985). We focus specifically on contexts in which episodic emotional support is delivered remotely by a nonrelational partner—conditions under which initial expectations about a support giver's experiential capacity, rather than nonverbal cues or relational history, are most likely to drive the support experience. This context also reflects a growing reality—from crisis hotlines and peer support platforms to AI chatbots, people are increasingly accessing remote support from nonrelational partners when close relational ties are unavailable. The present research specifically examines how initial expectations about AI's experiential capacity shape the experience and perception of AI support.

A growing body of research has begun documenting AI's potential in social support contexts. De Freitas et al. (2025) found that AI companion apps alleviate loneliness on par with interacting with another person, with users underestimating the degree to which these interactions help. Yin et al. (2024) showed that AI-generated responses made recipients feel more heard than human-generated responses. However, these same studies reveal that people's psychological responses to AI support are not simply a function of its quality. Yin et al. found that when recipients learned a response came from AI, they felt less heard, even when the response was identical. Furthermore, Dang and Liu (2024) showed that people deny humanness, especially emotional capacity, to those who even seek AI advice.

These findings converge on an important pattern: AI has the potential to produce functionally effective support, but people's beliefs about AI shape whether that support is experienced and evaluated positively. We propose that people's attribution of low experiential capacity to AI (compared with humans) and the expectations this attribution creates in social support contexts can help explain these patterns. In the following section, we develop this argument and derive two predictions about how initial expectations about AI's experiential capacity shape both how people experience and perceive AI support compared with human support.

How Initial Expectations About AI's Experiential Capacity Shape Support Experiences and Perceptions

Mind perception theory proposes that people attribute mind to entities along two fundamental dimensions: agency (the capacity to think, plan, and act) and experience (the capacity to sense, feel, and have emotional motivations; Epley & Waytz, 2010; H. M. Gray et al., 2007). Research on how people perceive AI's agency presents a nuanced picture. While early work suggested people attribute lower agency to machines than to humans (H. M. Gray et al., 2007), more recent findings indicate that people sometimes view AI as highly competent, particularly in domains involving analytical processing and decision making (Karataş & Cutright, 2023; Malle et al., 2025; Spatola & Urbanska, 2020), and may even hold AI to higher standards of accuracy than humans (Qin et al., 2025). What remains more consistent across the literature is that people perceive AI to have markedly lower experiential capacity than humans: a lower capacity to experience emotions and to have social motivations (Bigman et al., 2023; Bigman & Gray, 2018; H. M. Gray et al., 2007; Waytz et al., 2010). We propose that in social support contexts, it is people's initial expectations about this experiential dimension—the lay beliefs they bring to early encounters with AI—that shape how they experience and perceive AI support. This builds on prior work showing that people's expectations about interaction partners influence whether they seek or offer social support in human-to-human contexts (Dungan et al., 2022; Epley & Schroeder, 2014; Kardas et al., 2022; Kumar & Epley, 2023).

In the context of AI, we argue that expectations about experiential capacity are particularly consequential because they are closely intertwined with expectations about social motivation—not just what people think an entity is capable of but what they think it is motivated to do. As humans are perceived as having a full mind, capable of both evaluating someone and being socially motivated in that evaluation, their behavior is more readily attributed to motivations such as prejudice, care, or social judgment (Bigman et al., 2023). However, because AI is perceived as lacking experiential capacity, people are less likely to attribute social motivation to its behavior (Bigman et al., 2023; Bigman & Gray, 2018). In the context of receiving social support, we argue that this perceived absence of social motivation reduces FNE when receiving AI support. By contrast, when receiving support from another human, people may worry that the support giver will think poorly of them, blame them for their situation, or hold their disclosure against them. These concerns are rooted in the perception that a human support giver is socially motivated in the interaction—motivated to form impressions and evaluations and potentially act on those assessments. AI, perceived as lacking this social motivation, is unlikely to produce evaluations that carry such weight.

This reasoning is consistent with findings that people are more willing to disclose sensitive information to a virtual versus real human interviewer (Lucas et al., 2014) and are more open to their personal behaviors being tracked by algorithms than by humans (Raveendhran & Fast, 2021, 2024). In these cases, the reduced threat of evaluation appears to stem not from AI's inability to process information only but from perceptions that AI lacks the experiential capacity and the social motivation to be invested in the outcomes. Lowering FNE allows people to receive support without worrying about how they are perceived by the support giver, alleviating social stress and improving well-being (Crocker & Canevello, 2008; Gleason et al., 2008). We therefore predict that people will experience lower FNE when receiving support from AI (vs. humans), leading to improved well-being after minor aversive experiences.

The same initial expectations about AI's low experiential capacity also shape how people perceive the effectiveness of AI support. Empathy—the capacity to sense, share, and respond to another's emotions—is a key manifestation of experience (H. M. Gray et al., 2007) and a central ingredient in effective social support (Dakof & Taylor, 1990). People perceive AI as having lower experiential capacity and therefore assume it to be emotionally neutral (Jago & Laurin, 2022) and set low baseline expectations for its empathic responsiveness. By contrast, people expect human support givers to draw on their own emotional experiences in providing support (Ruttan et al., 2015) and to respond with genuine emotional depth (Dakof & Taylor, 1990). Paradoxically, this asymmetry in expectations creates an opportunity for AI: When AI systems produce empathic responses (Rubin et al., 2025), they are more likely to exceed people's modest initial expectations for empathy from AI, whereas humans face the task of surpassing already-high baselines. We therefore predict that people will perceive AI (vs. human) support as exceeding their initial expectations for empathy, which will in turn influence perceptions of support effectiveness.

In summary, both predictions stem from a common source: people's initial expectations about AI's lower experiential capacity compared with humans. People perceive AI as lacking the capacity to experience emotions and be socially motivated. As a result, they are less likely to fear its judgment, leading to improved well-being. They are also more likely to find that its empathy exceeds their expectations, leading to perceptions of support effectiveness. These two pathways illustrate how the same set of initial expectations about AI's experiential capacity can shape how AI support is experienced and perceived.

Overview of Studies

We begin with a preregistered pilot study showing that people hold different initial expectations about AI versus humans regarding their experiential capacity—specifically, propensity for negative social evaluation and empathy. We then test our main predictions across six studies (five preregistered). In Studies 1 and 2, participants report improved well-being when they believe support comes from AI rather than a human—an effect driven by reduced FNE that holds whether the human support giver is a coworker or a stranger. Study 3 extends this by showing that real interactions with AI (vs. humans) also lower FNE and improve well-being. Study 4 shows that although people initially expect AI to be less empathic than humans, real interactions with AI lead them to see it as surpassing those expectations more than human support givers do. Study 5 demonstrates that these exceeded expectations enhance perceived support effectiveness and examines alternative mechanisms. Study 6 replicates this finding in real interactions with both AI and human support givers. Together, these studies demonstrate that initial expectations about AI's experiential capacity shape how people experience and perceive AI support.

Across studies, we varied support delivery (audio, text), the human support giver's relational status (coworker, stranger), and the aversive encounters that prompted the need for support. In addition to simulation studies (Studies 1–2), four real-interaction studies (Studies 3–6) had participants discuss actual aversive experiences

with support givers—contexts in which support was delivered remotely by a nonrelational human partner or AI, consistent with the scope of our theoretical framework. We report all manipulations, exclusions, and measures and predetermined sample sizes to ensure 80% power to detect small to medium effects. For all real-interaction studies (Studies 3–6), prerecruitment filters were applied to exclude participants from prior studies. This research was approved by the first author's institutional review board. All data, preregistrations, and materials are posted on the Open Science Framework at https://tinyurl.com/ms4k35ns.

Pilot Study

We began by examining a foundational assumption of our model: that, consistent with mind perception theory (Epley & Waytz, 2010), people generally expect AI to have lower experiential capacity than humans. Specifically, we tested whether people assume that AI (vs. humans) has a lower capacity for socially motivated evaluations and empathy.

Method

Participants

We recruited 199 employed U.S. adults (gender: 97 women, 102 men; age: $M_{age} = 37.52$ years, $SD_{age} = 11.94$; ethnicity:[1] 23 African American, 20 Asian/Asian American, 152 White, 14 Hispanic, two Native American) through Prolific Academic in February 2025. Participants were paid $1.00 for the study (median completion time: approximately 3 min). We collected basic demographic information from our participants at the end of the study; responses were voluntary. This was a within-subjects experiment. A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 199 participants, $\alpha = .05$ (two-tailed), and 80% power, this within-subjects design could detect effect sizes as small as $dz = 0.20$. There were no exclusions. We preregistered this study on AsPredicted.org (see https://aspredicted.org/gvby-gssn.pdf).

Procedure

In this study, participants indicated their general expectations about humans' and AI's propensity for negative social evaluation and capacity for empathy. For both sets of measures, all items began with "In general, I expect humans [AI systems] to have the capacity to" and were rated on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). The order of the human and AI scales was counterbalanced.

Participants first rated their expectations about negative social evaluation on a three-item scale adapted from Leary's (1983) study ($\alpha_{human} = .93$; $\alpha_{AI} = .95$). The three items were "negatively judge humans in a given situation," "disapprove humans' behaviors in a given situation," and "form negative impressions of humans in a given situation." Then, participants rated their expectations about empathy on a three-item scale adapted from Davis's (1983) study ($\alpha_{human} = .94$; $\alpha_{AI} = .92$). The three items were "be concerned about humans in a given situation," "feel sorry for humans in a given situation," and "understand how humans feel in a given situation."

Results

General Expectations of AI's (vs. Humans') Propensity for Negative Social Evaluation

Consistent with our preregistration, we conducted a paired-samples *t* test to examine whether there were differences in participants' general expectations about humans' versus AI's propensity for negative social evaluation. Consistent with our prediction, we found significant differences between conditions such that participants, in general, expected that humans had a higher propensity to form negative social judgments ($M_{Humans} = 6.20$, $SD = 1.03$) compared with AI ($M_{AI} = 3.76$, $SD = 1.66$), $t(198) = 19.70$, $p < .001$, 95% CI of the difference [2.19, 2.68], $d = 1.40$, 95% CI [1.20, 1.59].

General Expectations of AI's (vs. Humans') Capacity for Empathy

Consistent with our preregistration, we conducted a paired-samples *t* test to examine whether there were differences in participants' general expectations about humans' versus AI's capacity for empathy. Consistent with our prediction, we found significant differences between conditions such that participants, in general, expected humans to have a higher capacity for empathy ($M_{Humans} = 6.41$, $SD = 0.84$) compared with AI ($M_{AI} = 2.69$, $SD = 1.59$), $t(198) = 26.30$, $p < .001$, 95% CI of the difference [3.44, 4.00], $d = 1.86$, 95% CI [1.63, 2.09].

In sum, these findings demonstrate that people hold markedly different initial expectations about AI versus humans: People expect AI to have a lower propensity for negative social evaluation and a lower capacity for empathy than humans. We argue that these initial expectations shape how people experience and perceive social support from AI (vs. humans) after minor aversive encounters.

Study 1

Study 1 examined whether receiving support from AI (vs. a human) after a minor aversive encounter would improve well-being and whether this effect would be mediated by reduced FNE.

Method

Participants

We recruited 600 employed U.S. adults (gender: 433 women, 167 men; age: $M_{age} = 29.67$ years, $SD_{age} = 9.34$; ethnicity: 48 African American, 54 Asian/Asian American, 470 White, 42 Hispanic, five Arab/Arab American, five Native American, seven other) through Prolific Academic in October 2021. Participants were paid $1.00 (median completion time: approximately 9 min). We collected basic demographic information from participants at the end of the study. This was a between-subjects experiment in which participants were randomly assigned to either a human support ($n = 302$) or AI support ($n = 298$) condition. A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 600 participants, $\alpha = .05$ (two-tailed), and 80% power, this between-subjects design could detect effect sizes as small as $d = 0.23$. There were no exclusions. We preregistered this study on AsPredicted.org (see https://aspredicted.org/blind.php?x=/KM5_L9P).

[1] For all studies, participants checked all ethnicity categories that applied.

Procedure

In this study, participants took part in a simulation in which they had a rude encounter with their supervisor. Participants were informed that they were employees in a university bookstore and were stuck in a major traffic jam due to a road construction project for 45 min, making them 30 min late for work. Following this, we asked participants to listen to recordings of what their supervisor had to say when they arrived late. In this recording, all participants heard the supervisor talk to them rudely when they tried to explain why they were late. Following that aversive experience, participants were asked to describe what happened to a support giver and received support purportedly either from a human or an AI depending on the condition to which they were assigned. Participants in both conditions were asked to write about their interaction with their supervisor and how it made them feel and share these thoughts with the support giver.

Next, we manipulated the support provider. Participants in the human support condition were informed that they were sending a chat message about this interaction to their coworker with whom they typically discussed work situations. Participants in the AI support condition were informed that they were sending a chat message to an AI system. In both conditions, participants were told that the support provider (either the human or AI) would offer support to try to make them feel better.

After participants wrote about the interaction, they were informed that their coworker or the AI system would read their message and offer support. To increase realism in the simulation, participants in both conditions were asked to wait for a short period of time while the support provider purportedly formulated a response and then were informed that the support provider was ready for them. Following this, they listened to an audio recording of the support in both conditions. Importantly, participants in both the human and AI support conditions received the exact same support message.

Following this, all participants rated their FNE and three well-being outcomes: negative affect, anxiety, and stress. Drawing on research demonstrating that these represent related but distinct negative emotional states (Lovibond & Lovibond, 1995), we assessed them separately to assess well-being: Negative affect captures general emotional distress, anxiety reflects apprehensive arousal and worry, and stress reflects perceived strain and tension. Full manipulations are included in the Open Science Framework repository.

Measures

Participants rated items on a scale ranging from 1(*not at all*) to 5 (*very much*).

FNE. We measured FNE using an eight-item scale developed by Leary (1983; $\alpha = .90$). The items included the following: "I feel worried about being judged by [support giver]," "I feel that [the support giver] noticed my shortcomings," "I am concerned that my behaviors were not approved of by [support giver]," "I am concerned that [support giver] found faults with me," "I am concerned about what [support giver] thought about me," "I am worried about the kind of impression I made on [support giver]," "I am concerned about how I came across to [support giver]," and "I am concerned that I said the wrong things to [support giver]."

Negative Affect. We measured participants' negative affect using the 10 negative emotion items from the Job-Related Affective Well-being Scale (Van Katwyk et al., 2000), ($\alpha = .89$). Items included "annoyed," "frustrated," "discouraged," "angry," "confused," "furious," "gloomy," "anxious," "frightened," and "intimidated."[2]

Anxiety. We measured anxiety using a four-item scale developed by Marteau and Bekker (1992; $\alpha = .80$). Items included the following: "I feel calm" (reverse coded), "I feel tense," "I feel nervous," "I feel pleasant" (reverse coded).

Stress. We measured stress using a four-item scale developed by Stanton et al. (2001; $\alpha = .87$). Items included the following: "I feel pressured," "I feel stressed," "I feel irritated," and "I feel overwhelmed."

Results

FNE

Results of an independent-samples t test revealed significant differences between conditions. Consistent with our prediction, participants in the AI support condition reported feeling lower FNE ($M_{\text{AI Support}} = 3.00$, $SD = 1.04$) than those in the human support condition ($M_{\text{Human Support}} = 3.45$, $SD = 0.99$), $t(598) = -5.44$, $p < .001$, 95% CI of the difference [−.61, −.29], $d = -.44$, 95% CI [−.61, −.28].

Negative Affect

Participants in the AI support condition reported lower feelings of negative affect ($M_{\text{AI Support}} = 2.80$, $SD = 0.97$) than those in the human support condition ($M_{\text{Human Support}} = 3.06$, $SD = 0.85$), $t(598) = -3.46$, $p = .001$, 95% CI of the difference [−.40, −.11], $d = -.28$, 95% CI [−.44, −.12].

Anxiety

Participants in the AI support condition reported lower anxiety ($M_{\text{AI Support}} = 3.19$, $SD = 1.05$) than those in the human support condition ($M_{\text{Human Support}} = 3.73$, $SD = 0.87$), $t(598) = -6.92$, $p < .001$, 95% CI of the difference [−.70, −.39], $d = -.57$, 95% CI [−.73, −.40].

Stress

Participants in the AI support condition reported lower stress ($M_{\text{AI Support}} = 2.97$, $SD = 1.23$) than those in the human support condition ($M_{\text{Human Support}} = 3.39$, $SD = 1.05$), $t(598) = -4.50$, $p < .001$, 95% CI of the difference [−.60, −.24], $d = -.37$, 95% CI [−.53, −.21] (see Figure 1).

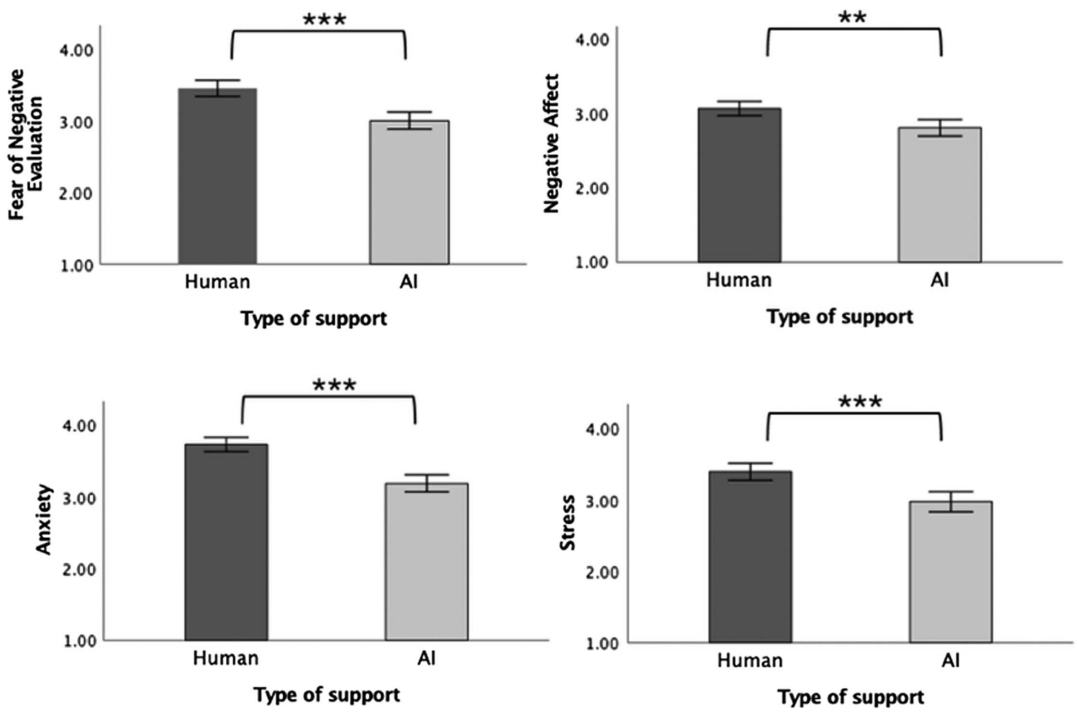
Mediation Analyses

To test our prediction that FNE mediated the effect of AI (vs. human) support on our well-being outcomes—negative affect,

[2] Our preregistration for Studies 1 and 2 also included the Positive and Negative Affect Schedule negative affect scale (Watson & Clark, 1994) as a separate measure (Study 1: $\alpha = .88$; Study 2: $\alpha = .85$). Because subsequent studies used only the focal negative affect measure based on the Job-Related Affective Well-being Scale, we report that measure across all studies for consistency. All results replicate using the Positive and Negative Affect Schedule negative affect scale (see Supplemental Materials).

Figure 1

Participants Reported Lower Fear of Negative Evaluation, Lower Negative Affect, Lower Anxiety, and Lower Stress After AI Support (Study 1)



Note. Error bars represent 95% confidence interval.
** $p < .01$. *** $p < .001$.

anxiety, and stress—we conducted a series of bootstrapping mediation analyses using the PROCESS macro (Model 4; Hayes, 2013) with FNE as the mediator and each of the well-being outcomes as the dependent variable. We conducted a bootstrapping analysis based on a resampling size of 10,000 for each of the dependent variables. Results revealed that the 95% bias-corrected confidence interval for the indirect effect of AI support through FNE on negative affect (95% CI [−.26, −.11]), anxiety (95% CI [−.23, −.10]), and stress (95% CI [−.30, −.13]) did not include zero, indicating that FNE mediated the relationship between AI support and each of the well-being outcomes (See Figure 2).

Study 2

In Study 1, the human support giver was a coworker—someone familiar—which may have heightened FNE in the human condition. In Study 2, we added a third condition in which participants received support from a human stranger to rule out the possibility that the effects observed in Study 1 were influenced by the support giver's familiarity.

Method

Participants

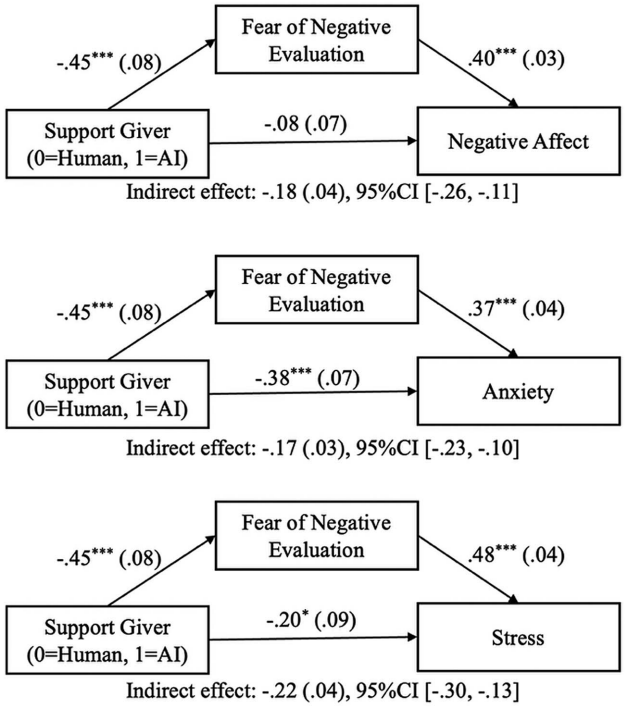
We recruited 1,099 employed U.S. adults (gender: 540 women, 558 men, one missing; age: M_{age} = 34.52 years, SD_{age} = 10.68; ethnicity: 97 African American, 135 Asian/Asian American, 807 White, 90 Hispanic, six Arab/Arab American, 16 Native American, 13 other) through Prolific Academic in January 2021. Participants were paid $1.00 (median completion time: approximately 9 min). We collected basic demographic information at the end of the study. This was a between-subjects experiment in which participants were assigned to one of three conditions: human support stranger (n = 365), human support coworker (n = 367), and AI support (n = 367). A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 1,099 participants, $\alpha = .05$, and 80% power, this between-subjects design could detect effect sizes as small as $f = 0.09$. There were no exclusions. We preregistered this study on AsPredicted.org (see https://aspredicted.org/LYE_XRE).

Procedure

The procedure was identical to Study 1, with the addition of the third condition. After the aversive encounter, participants in all three conditions received the same support message via an audio recording as in Study 1. Only the described source of the support varied: Participants were informed that the support came from a coworker (human coworker condition), a person who worked at a support hotline (human stranger condition), or an AI system (AI condition). Full manipulations are included in the Open Science Framework repository.

Figure 2

Fear of Negative Evaluation Mediates the Effect of AI Support (vs. Human Support) on Well-Being (Negative Affect, Anxiety, Stress; Study 1)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
$^{*}p < .05$. $^{***}p < .001$.

Measures

Participants rated items on a scale ranging from 1 (*not at all*) to 5 (*very much*). We measured FNE ($\alpha = .92$), negative affect ($\alpha = .90$), anxiety ($\alpha = .78$), and stress ($\alpha = .87$) using the same scales as in Study 1 (see Footnote 1).

Results

FNE

Results of a one-way ANOVA revealed significant differences in FNE between the three conditions, $F(2, 1096) = 14.69$, $p < .001$, $\eta^2 = .03$, 95% CI [.01, .05]. As expected, participants in the AI support condition ($M_{\text{AI support}} = 2.80$, $SD = 1.11$) reported lower levels of FNE than those in the human support coworker ($M_{\text{human-coworker}} = 3.21$, $SD = 1.04$; $p < .001$) and human support stranger ($M_{\text{human-stranger}} = 3.12$, $SD = 1.08$; $p < .001$) conditions. There were no significant differences between the two human conditions ($p = .29$).

Negative Affect

Results of a one-way ANOVA indicated significant differences among the three conditions, $F(2, 1095) = 6.17$, $p = .002$, $\eta^2 = .01$, 95% CI [.002, .03]. Participants in the AI support condition ($M_{\text{AI support}} = 2.71$, $SD = 0.99$) reported lower levels of negative affect than those in the human support coworker ($M_{\text{human-coworker}} = 2.92$, $SD = 0.94$; $p = .003$) and in the human support stranger ($M_{\text{human-stranger}} = 2.93$, $SD = 0.94$; $p = .002$) conditions, and there were no significant differences between the two human conditions ($p = .87$).

Anxiety

A one-way ANOVA revealed significant differences between the three conditions, $F(2, 1096) = 12.18$, $p < .001$, $\eta^2 = .02$, 95% CI [.007, .04]. Participants in the AI support condition ($M_{\text{AI support}} = 3.26$, $SD = 0.95$) reported lower levels of anxiety than those in the human support coworker ($M_{\text{human-coworker}} = 3.59$, $SD = 0.91$; $p < .001$) and in the human support stranger ($M_{\text{human-stranger}} = 3.50$, $SD = 0.97$; $p < .001$) conditions, and there were no differences between the two human conditions ($p = .19$).

Stress

A one-way ANOVA revealed significant differences between the three conditions, $F(2, 1096) = 8.76$, $p < .001$, $\eta^2 = .02$, 95% CI [.004, .03]. Participants in the AI support condition ($M_{\text{AI support}} = 2.95$, $SD = 1.19$) reported lower levels of stress than those in the human support coworker ($M_{\text{human-coworker}} = 3.25$, $SD = 1.09$; $p < .001$) and human support stranger ($M_{\text{human-stranger}} = 3.25$, $SD = 1.11$; $p < .001$) conditions, and there were no differences between the two human conditions ($p = .92$; see Figure 3).

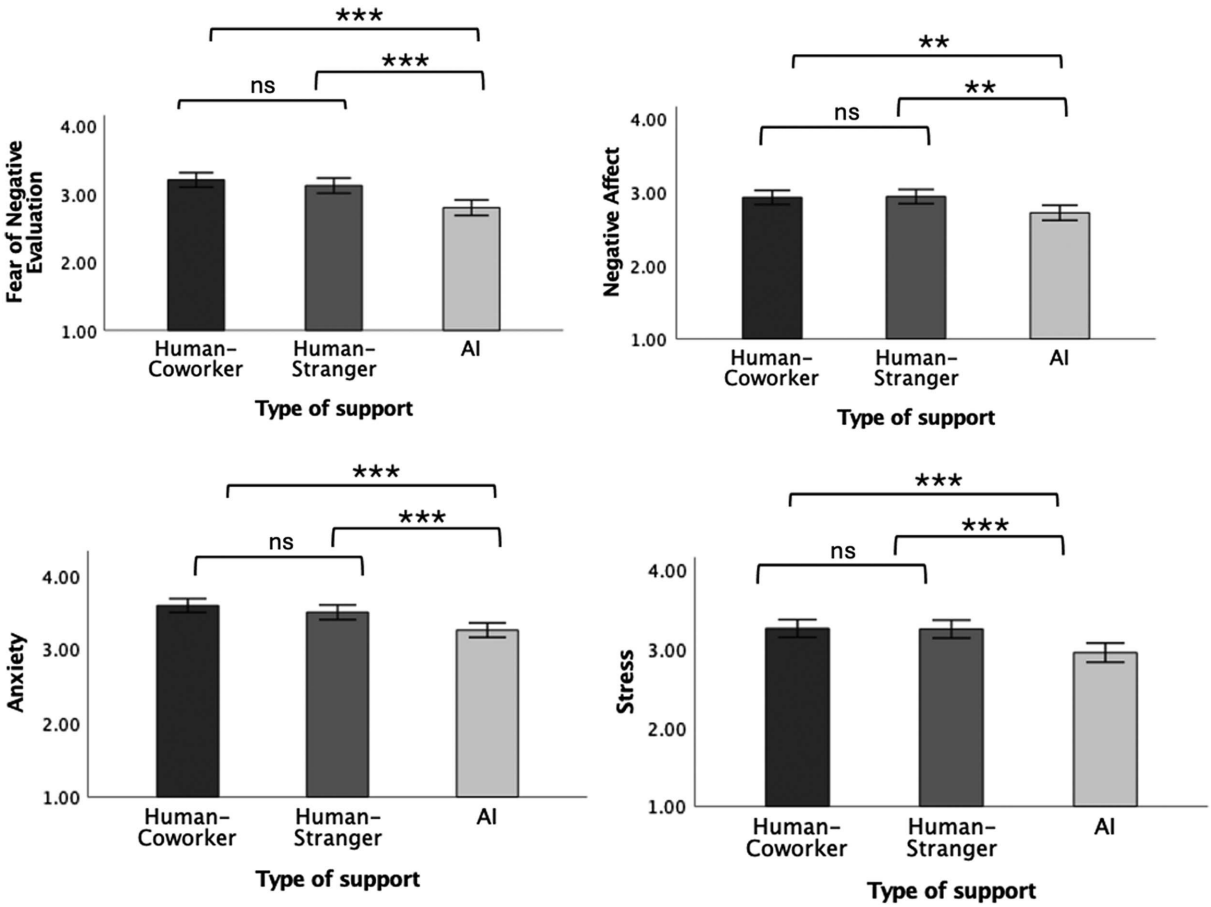
Mediation Analyses

Given the categorical nature of our independent variable (human coworker, human stranger, and AI conditions), we used the indicator coding procedure described by Hayes and Preacher (2013) with 10,000 bootstrap resamples to calculate the indirect effects of AI support on our well-being outcomes via FNE. The AI condition was specified as the reference group. For interpretive clarity, we reverse coded the resulting contrasts such that negative coefficients indicate lower anxiety in the AI condition relative to the human conditions. When compared with the human stranger condition, AI support had a significant indirect effect via FNE on negative affect (95% CI [−.22, −.07]), anxiety (95% CI [−.19, −.06]), and stress (95% CI [−.26, −.09]). Similarly, when compared with the human coworker condition, AI support had a significant indirect effect via FNE on negative affect (95% CI [−.26, −.12]), anxiety (95% CI [−.22, −.09]), and stress (95% CI [−.31, −.14]; see Figure 4).

Consistent with our preregistered analysis plan, we also combined the two human conditions (human stranger and human coworker) and compared this combined condition with the AI condition by conducting independent-samples *t* tests. Consistent with our predictions and findings from Study 1, participants in the AI support condition reported feeling lower FNE than those in the combined human support condition, $M_{\text{AI Support}} = 2.80$, $SD = 1.11$; $M_{\text{Human Support}} = 3.17$, $SD = 1.06$, $t(1097) = -5.32$, $p < .001$, 95% CI of the difference [−.50, −.23], $d = -.34$, 95% CI [−.47, −.21]. Furthermore, participants experienced improved well-being outcomes in the AI support condition than this combined human support condition across all indicators of well-being (i.e., lower negative affect, lower anxiety, and lower stress). Specifically, participants in the AI support condition reported feeling lower negative affect, $M_{\text{AI Support}} = 2.71$, $SD = 0.99$; $M_{\text{Human Support}} = 2.93$, $SD = 0.94$, $t(1096) = -3.51$, $p < .001$,

Figure 3

Participants Reported Lower Fear of Negative Evaluation, Lower Negative Affect, Lower Anxiety, and Lower Stress After AI Support Relative to Human (Stranger and Coworker) Support (Study 2)



Note. Error bars represent 95% confidence interval. *ns* = nonsignificant.
** $p < .01$. *** $p < .001$.

95% CI of the difference [−.34, −.09], $d = -.23$, 95% CI [−.35, −.10]; lower anxiety, $M_{\text{AI Support}} = 3.26$, $SD = 0.95$; $M_{\text{Human Support}} = 3.55$, $SD = 0.94$, $t(1097) = -4.76$, $p < .001$, 95% CI of the difference [−.41, −.17], $d = -.30$, 95% CI [−.43, −.18]; and lower stress, $M_{\text{AI Support}} = 2.95$, $SD = 1.19$; $M_{\text{Human Support}} = 3.25$, $SD = 1.10$, $t(1097) = -4.19$, $p < .001$, 95% CI of the difference [−.44, −.16], $d = -.27$, 95% CI [−.39, −14], than those in the combined human support condition.

We also examined whether FNE mediated the effect of AI support compared with the combined human conditions on the well-being measures by conducting a series of bootstrapping mediation analyses based on a resampling size of 10,000 using the PROCESS macro (Model 4; Hayes, 2013) with FNE as the mediator and each of the well-being outcomes as dependent variables. Results revealed that the 95% bias-corrected confidence interval for the indirect effect of AI support through FNE on negative affect (95% CI [−.23, −.10]), anxiety (95% CI [−.19, −.09]), and stress (95% CI [−.27, −.12]) did not include zero, indicating that FNE mediated the relationship between AI support and each of the well-being outcomes (see Figure 5).

In sum, these results suggest that receiving support from AI elicits lower FNE than receiving support from a human—whether a coworker or a stranger—which, in turn, is associated with better well-being outcomes.

Study 3

Studies 1 and 2 provided evidence that AI (vs. human) support after minor aversive encounters evokes lower FNE, which is associated with better well-being outcomes. However, both studies used simulated support interactions with scripted responses. In Study 3, participants described their own aversive experience and received support in real time from a human or AI support giver, providing a more ecologically valid test of our predictions.

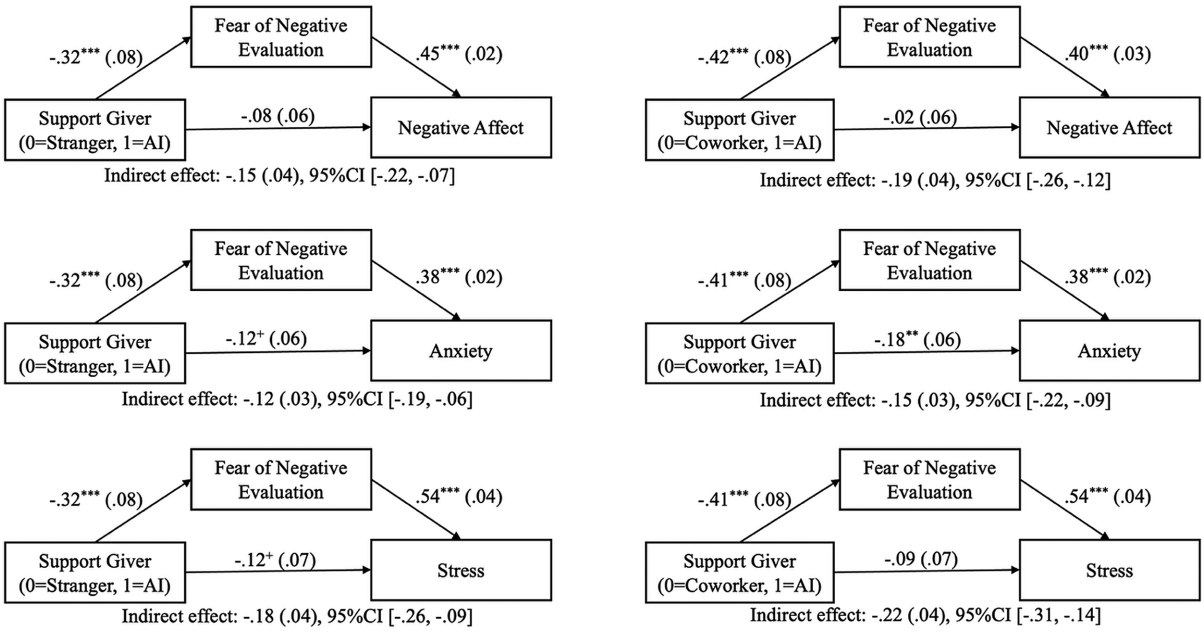
Method

Participants

We recruited 448 employed U.S. adults (gender: 196 women, 245 men, seven other; age: $M_{\text{age}} = 39.10$ years, $SD_{\text{age}} = 12.88$; ethnicity: 39 African American, 47 Asian/Asian American, 283 White, 33 Hispanic, one Arab/Arab American, five other) through Prolific

Figure 4

Fear of Negative Evaluation Mediates the Effect of AI Support (vs. Human Support) on Well-Being (Negative Affect, Anxiety, Stress; Study 2)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
$^{\dagger}p < .10$. $^{**}p < .01$. $^{***}p < .001$.

Academic in July to August 2023. In the human support condition, we recruited an additional 222 (gender: 85 women, 136 men; age: M_{age} = 39.79 years, SD_{age} = 12.48) individuals to serve as support givers who interacted in real time with participants; no measures were collected from these individuals.[3] Participants were paid between $2.00 and $2.60 for the study (median completion time: approximately 9 min). We collected basic demographic information from our participants at the end of the study. This was a between-subjects experiment in which participants were randomly assigned to either a human support (n = 222) or AI support (n = 226) condition. A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 448 participants, $\alpha = .05$ (two-tailed), and 80% power, this between-subjects design could detect effect sizes as small as $d = 0.27$. There were no exclusions. We preregistered this study on AsPredicted.org (see https://aspredicted.org/9NX_J8F).

Procedure

In this study, participants interacted in real time[4] with the support giver—human or AI depending on the condition they were randomly assigned to. All participants were first asked to think about the last time they had a bad encounter with another person and describe what happened and how they felt. Participants described a range of minor aversive social experiences, such as being treated rudely by a stranger at a restaurant, being cut off while driving, dealing with a disrespectful neighbor, or having a disagreement with a coworker. They then sent this description to their assigned support giver via the chat interface.

AI Condition. To enable real-time interactions between participants and the AI, we created a customized AI chatbot interface using the OpenAI software development kit powered by GPT-3.5 Turbo (https://openai.com/). This custom interface allowed us to preset the role and the goal for the AI while leveraging a ChatGPT-like experience. This AI model acted as a support giver and provided supportive responses to participants to help them feel better about their situation.

Human Condition. To enable real-time interactions between participants and a human support giver, we used ChatPlat, an online application validated in prior research that allows researchers to design chat interfaces wherein people can interact in real time (e.g., Huang et al., 2017; Wolf et al., 2016). Participants were informed that they would interact with another person participating in this study and were randomly paired with a person recruited to play the role of the support giver. The human support giver was instructed to provide support to help them feel better about their situation.

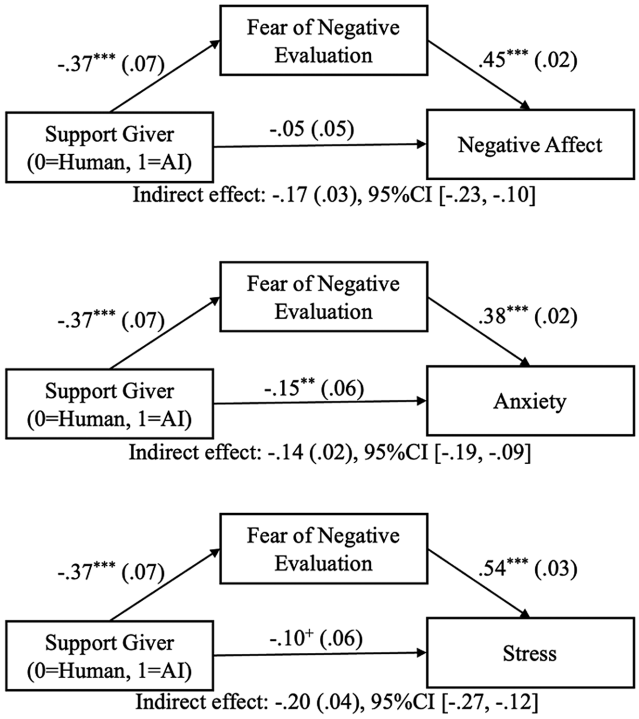
In both conditions, we customized the chat interface such that each participant would be able to send only one message to the support giver and the support giver would also be able to send only one message as a reply. This allowed us to control for conversation length within and between the experimental conditions. After receiving the support message, participants in both conditions provided their ratings on FNE and the three well-being outcomes—negative affect, anxiety, and stress—used in previous studies.

[3] In the human condition, we overrecruited to account for issues with real-time pairings.

[4] Across all real interaction studies (Studies 3–6), we conducted post hoc data quality checks to identify participants who self-reported technical difficulties. Full details, including exclusion rates and robustness analyses, are reported in the Supplemental Materials.

Figure 5

Fear of Negative Evaluation Mediates the Effect of AI Support (vs. Human Support) on Well-Being (Negative Affect, Anxiety, Stress; Study 2—Human Stranger and Coworker Combined)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
$^{\dagger}p < .10$. $^{**}p < .01$. $^{***}p < .001$.

Measures

Unless otherwise noted, participants rated items on a scale ranging from 1 (*not at all*) to 5 (*very much*). We measured FNE ($\alpha = .94$), negative affect ($\alpha = .91$), anxiety ($\alpha = .77$), and stress ($\alpha = .85$) using the same scales as in previous studies.

Results

As preregistered, we examined (a) whether there were differences in FNE ratings between the AI and the human conditions and (b) whether FNE mediated the effect of AI (vs. human) support on the well-being outcomes. Consistent with previous studies, we expected (and preregistered) an indirect effect of AI support on our well-being outcomes, via FNE. Unlike Studies 1 and 2, where all participants received identical support in both conditions, in this study each participant received unique support from their partner, introducing substantial variance in the quality of the support participants received. We therefore preregistered the indirect effect through FNE but did not expect (nor preregister) a main effect of support source on the well-being measures.[5]

FNE

Consistent with our preregistration, we first examined whether support source influenced FNE. Results of an independent-samples *t* test revealed significant differences between conditions. Consistent with our prediction, participants in the AI support condition reported feeling lower FNE ($M_{\text{AI Support}} = 1.39$, $SD = 0.62$, 95% CI [1.31, 1.48]) than those in the human support condition ($M_{\text{Human Support}} = 1.74$, $SD = 0.96$, 95% CI [1.61, 1.87]), $t(446) = -4.53$, $p < .001$, 95% CI of the difference [−.49, −.20], $d = -.43$, 95% CI [−.61, −.24].

Mediation Analyses

Next, consistent with our preregistration, we explored whether FNE mediated the effect of support source on well-being. To examine this relationship, we conducted a series of bootstrapping mediation analyses using the PROCESS macro (Model 4; Hayes, 2013) with FNE as the mediator and each of the well-being outcomes as dependent variables. We conducted a bootstrapping analysis based on a resampling size of 10,000 for each of the dependent variables.

Consistent with our prediction and findings from previous studies, results revealed that the 95% bias-corrected confidence interval for the indirect effect of AI support through FNE on negative affect (95% CI [−.17, −.04]), anxiety (95% CI [−.18, −.05]), and stress (95% CI [−.18, −.05]) did not include zero, indicating that FNE mediated the relationship between AI support and each of the well-being measures (see Figure 6). In sum, our findings revealed that when having real interactions with an AI (vs. human) support giver, people experienced lower FNE, and this, in turn, mediated the effect of AI support on negative affect, anxiety, and stress.

Textual Analysis

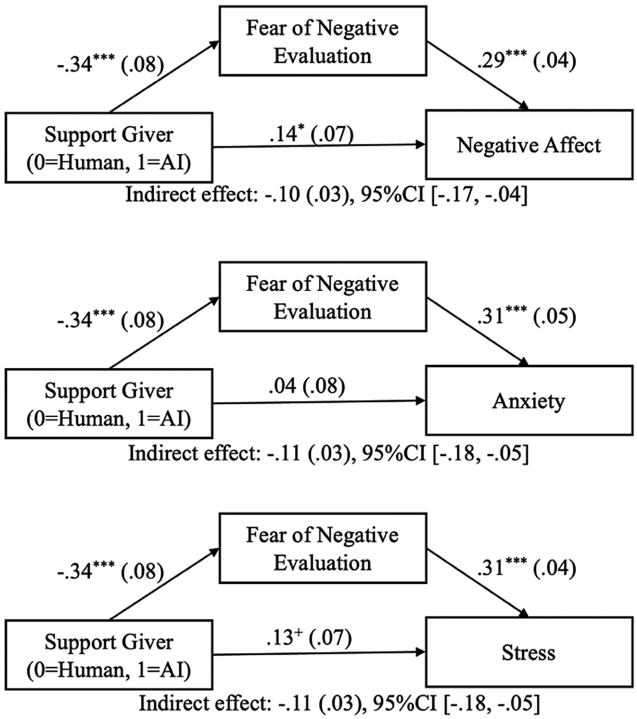
It is possible that participants disclosed different types of information to AI versus human support givers that could influence how they experienced the support interaction. To explore this, we conducted an exploratory textual analysis of participants' interactions to examine whether there were systematic differences in the content that participants shared across the two conditions. To do so, we leveraged the Linguistic Inquiry and Word Count method (LIWC; Pennebaker et al., 2015), which uses prevalidated dictionaries to quantitatively analyze text (Kacewicz et al., 2014).

Because our goal was to examine whether participants described their negative experiences differently across conditions, we used LIWC dictionaries related to negative emotion, anxiety, anger, and sadness. In other words, we focused on whether participants expressed different levels of distress when writing to AI versus human support givers. Results indicated that there were no significant differences in any of these indicators: negative emotions, $M_{\text{human}} = 4.69$, $SD_{\text{human}} = 3.04$; $M_{\text{AI}} = 4.73$, $SD_{\text{AI}} = 3.48$; $F(1, 446) = .02$, $\eta^2 = .00$, 95% CI [.000, .007]; anxiety, $M_{\text{human}} = 1.15$, $SD_{\text{human}} = 1.39$; $M_{\text{AI}} = 1.28$, $SD_{\text{AI}} = 1.73$; $F(1, 446) = .76$, $p = .39$, $\eta^2 = .002$, 95% CI [.000, .018]; anger, $M_{\text{human}} = 1.61$, $SD_{\text{human}} = 1.80$; $M_{\text{AI}} = 1.56$, $SD_{\text{AI}} = 2.16$; $F(1, 446) = .07$, $p = .79$, $\eta^2 = .00$, 95% CI [.000, .010]; or sadness, $M_{\text{human}} = .53$, $SD_{\text{human}} = 1.07$; $M_{\text{AI}} = .50$, $SD_{\text{AI}} = 1.05$; $F(1, 446) = .06$, $p = .81$, $\eta^2 = .00$, 95% CI [.000, .009], between the two conditions. Thus, our findings suggest that

[5] We did not find significant differences between conditions on any of the well-being measures ($p = .45 - .73$).

Figure 6

Fear of Negative Evaluation Mediates the Effect of AI Support (vs. Human Support) on Well-Being (Negative Affect, Anxiety, Stress; Study 3)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
$^{+}p < .10$. $^{*}p < .05$. $^{***}p < .001$.

participants did not systematically differ in what they shared with AI versus human support givers.

Study 4

Studies 1–3 provided evidence that AI (vs. human) support after minor aversive encounters evokes lower FNE, which, in turn, is associated with better well-being outcomes. Study 4 turns to the second pathway in our model: perceptions of AI support. Specifically, we tested whether people's initial expectation that AI has a lower capacity for empathy than humans creates an opportunity for AI to exceed those empathy expectations more than human support givers.

Method

Participants

We recruited 499 employed U.S. adults (gender: 286 women, 213 men; age: $M_{age} = 37.24$ years, $SD_{age} = 11.94$; ethnicity: 93 African American, 39 Asian/Asian American, 336 White, 39 Hispanic, six Arab/Arab American, 10 Native American, six other) through Prolific Academic in November 2024. Participants were paid $2.00 (median completion time: approximately 9 min). We collected basic demographic information from participants at the end of the study. This was a between-subjects experiment in which participants were randomly assigned to either a human support ($n = 248$) or AI support ($n = 251$) condition. A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 499 participants, $\alpha = .05$ (two-tailed), and 80% power, this between-subjects design could detect effect sizes as small as $d = 0.25$. There were no exclusions. We preregistered this study on AsPredicted.org (see https://aspredicted.org/b3r8-r9t8.pdf).

Procedure

In this study, participants first indicated their general expectations about humans' and AI's capacity for empathy using the same three-item scale as in the pilot study (adapted from Davis, 1983; $\alpha_{human} = .91$; $\alpha_{AI} = .91$). All ratings were made on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Following this, participants were informed that they would interact with another person (human condition) or an AI system (AI condition) to receive support related to an aversive encounter they had and answer questions about this interaction.

Custom Chatbot for Real-Time Interaction. We built a custom chatbot interface using the OpenAI software development kit, leveraging the GPT-4o model, one of the most advanced AI models available at the time of this study that has been consistently rated by human evaluators as producing high-quality, coherent, and contextually relevant responses. We preset the role and the goal for the AI and also designed the chatbot to maintain sufficient memory throughout the session. This allowed participants to engage in multiple back-and-forth exchanges while ensuring that responses remained contextually grounded throughout the conversation. This AI model acted as a support giver and provided supportive responses to participants' input.

Real-Time Interactions With the Support Giver. In both conditions, all participants interacted in real time with the same AI chatbot. We experimentally manipulated their beliefs about who they were interacting with to isolate the effect of psychological expectations, independent of any actual differences in support quality. Participants in the human condition were informed that they were interacting with a person who worked at a support hotline and was trained to help them. In the AI condition, participants were informed that they were interacting with an AI chatbot that was trained to help them.

Participants in both conditions were asked to think about the last time they had a negative encounter with another person and describe what happened and how they felt. They sent this as a message via the chat interface we created to enable participants to interact with the support giver in real time. Participants then received a supportive response and were able to have back-and-forth exchanges with the support giver.

Following this, participants indicated the extent to which the support giver met their expectations for empathy on the following item: "Please indicate the extent to which [the support giver] met your expectations for empathy": 0 (less empathetic than I expected), 50 (as empathetic as I expected), and 100 (more empathetic than I expected).

Results

General Expectations of Capacity for Empathy

Consistent with our preregistration, we conducted a paired-samples t test to examine whether there were differences in

participants' general expectations about humans' versus AI's capacity for empathy prior to the real-time interaction with the support giver. Consistent with our predictions, results revealed significant differences such that participants expected AI to have lower capacity for empathy ($M_{AI} = 2.77$, $SD = 1.71$) compared with humans, $M_{Humans} = 6.18$, $SD = 1.03$, $t(498) = -35.83$, $p < .001$, 95% CI of the difference [−3.59, −3.22], $d = -1.60$, 95% CI [−1.74, −1.47].

Postsupport Perceptions About Empathy Expectations

Next, we examined whether AI exceeded participants' empathy expectations more than the human support giver after the real-time interaction. Consistent with our prediction, results of an independent-samples t test revealed that participants in the AI support condition indicated that AI exceeded their expectations for empathy ($M_{AI\ Support} = 71.92$, $SD = 21.76$) more than participants in the human support condition ($M_{Human\ Support} = 65.49$, $SD = 24.22$), $t(497) = 3.12$, $p = .002$, 95% CI of the difference [2.38, 10.48], $d = .28$, 95% CI [.10, .46].

Textual Analysis

Similar to Study 3, we explored whether participants disclosed different types of information to AI versus human support givers through an exploratory textual analysis of participants' interactions using LIWC (Pennebaker et al., 2015). Similar to Study 3, our goal was to examine whether participants expressed different levels of distress when writing to AI versus human support givers. Therefore, we used LIWC dictionaries related to negative emotion, anxiety, anger, and sadness. Results indicated that there were no significant differences between the two conditions in any of these indicators: negative emotions, $M_{human} = 4.84$, $SD_{human} = 5.01$; $M_{AI} = 4.14$, $SD_{AI} = 2.91$; $F(1, 497) = 3.71$, $p = .06$, $\eta^2 = .007$, 95% CI [.000, .029]; anxiety, $M_{human} = 1.06$, $SD_{human} = 1.29$; $M_{AI} = 1.01$, $SD_{AI} = 1.29$; $F(1, 497) = .19$, $p = .66$, $\eta^2 = .00$, 95% CI [.000, .011]; anger, $M_{human} = 1.94$, $SD_{human} = 4.61$; $M_{AI} = 1.40$, $SD_{AI} = 1.78$; $F(1, 497) = 3.02$, $p = .08$, $\eta^2 = .006$, 95% CI [.000, .027]; or sadness, $M_{human} = .54$, $SD_{human} = 1.06$; $M_{AI} = .52$, $SD_{AI} = 1.11$; $F(1, 497) = .06$, $p = .81$, $\eta^2 = .00$, 95% CI [.000, .008].

In sum, although people generally expect AI to have a lower capacity for empathy than humans, they perceived AI as exceeding those expectations to a greater degree than humans after receiving support. Notably, all participants interacted with the same chatbot, differing only in whether it was labeled as AI or human, highlighting the role of psychological expectations in shaping these perceptions.

Study 5

Study 4 established that people's initial expectation that AI has a lower capacity for empathy creates an opportunity for AI to exceed those expectations more than human support givers. Study 5 took the next step by testing whether this tendency for AI to exceed empathy expectations, in turn, drives greater perceived support effectiveness. In addition, we explored two alternative mechanisms that could also contribute to perceptions of AI support effectiveness: the belief that the support giver might help deal with future problems and the sense that one can express negative emotions more freely.[6]

Method

Participants

We recruited 294 employed U.S. adults (gender: 179 women, 115 men; age: $M_{age} = 40.88$ years, $SD_{age} = 17.93$; ethnicity: 64 African American, 16 Asian/Asian American, 201 White, 11 Hispanic, six Native American, three other) through Prolific Academic in September 2024. Participants were paid $2.00 for the study (median completion time: approximately 13 min). We collected basic demographic information from our participants at the end of the study. This was a between-subjects experiment in which participants were randomly assigned to either a human support ($n = 135$) or AI support ($n = 159$) condition. A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 294 participants, $\alpha = .05$ (two-tailed), and 80% power, this between-subjects design could detect effect sizes as small as $d = 0.33$. There were no exclusions.

Procedure

As in Study 4, participants in both conditions engaged in real-time interactions with the same custom chatbot interface developed using OpenAI's software development kit, leveraging GPT-4o. Participants in the human condition were informed that they were interacting with a person who worked at a support hotline who was trained to help them, while those in the AI condition were informed that they were interacting with an AI chatbot that was trained to help them. In both conditions, participants were asked to think about the last time they had a negative encounter with another person and describe what happened and how they felt. They sent this as a message via the chat interface, received a supportive response, and were able to engage in back-and-forth exchanges with the support giver. Following the interaction, participants indicated the extent to which the support giver met their expectations for empathy, rated their perceptions of support effectiveness, and provided ratings for two exploratory measures for alternative mechanisms.

Measures

Postsupport Perceptions About Empathy Expectations. Participants indicated the extent to which the support giver met their expectations for empathy using the same item as in Study 4: "Please indicate the extent to which [the support giver] met your expectations for empathy": 0 (less empathetic than I expected), 50 (as empathetic as I expected), and 100 (more empathetic than I expected).

Perceptions of Support Effectiveness. We measured perceptions of support effectiveness using a three-item scale developed for this study that reflected and matched the actual measures of well-being (i.e., anxiety, stress, negative affect) used in Studies 1–3, so that they appropriately reflected people's perceptions of the extent to which their interaction with the support giver (human vs. AI) resulted in reducing these specific outcomes after negative encounters. These items included the following: "I can really count on this form of support to help me feel less anxious after negative encounters,"

[6] This study was not preregistered; it was originally conducted as a supplemental study to provide an initial test of the empathy expectations pathway while also exploring alternative mechanisms identified during the review process.

"I can really count on this form of support to help me feel less stressed after negative encounters," and "I can really count on this form of support to help me improve my mood after negative encounters" (α = .97).

Exploratory Alternative Mechanisms. We measured two potential alternative mechanisms on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*): the possibility that the support giver might help deal with future negative encounters ("[Alex/This AI system] can help me deal with other issues in the future") and the possibility that one might express negative emotions more freely when interacting with the support giver ("When talking with [Alex/This AI system], I was able to express my negative emotions more freely").

Results

Postsupport Perceptions About Empathy Expectations

Consistent with Study 4, results of an independent-samples *t* test revealed that participants in the AI support condition reported that the support giver exceeded their expectations for empathy ($M_{\text{AI Support}}$ = 59.27, *SD* = 23.98) more than participants in the human support condition ($M_{\text{Human Support}}$ = 50.13, *SD* = 28.82), $t(292) = 2.97$, $p = .003$, 95% CI of the difference [3.07, 15.20], $d = .35$, 95% CI [.12, .58].

Perceptions of Support Effectiveness

An independent-samples *t* test revealed that participants in the AI support condition perceived the support they received as more effective ($M_{\text{AI Support}} = 3.00$, $SD = 1.18$) compared with participants in the human support condition ($M_{\text{Human Support}} = 2.59$, $SD = 1.33$), $t(292) = 2.86$, $p = .005$, 95% CI of the difference [.13, .71], $d = .33$, 95% CI [.10, .56].

Mediation Analyses

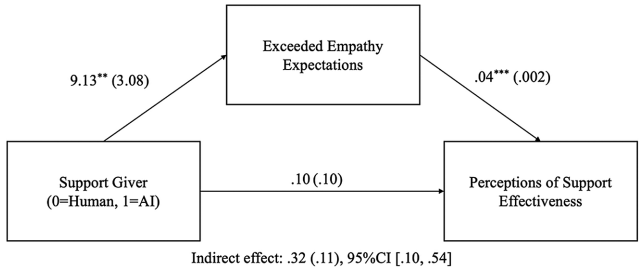
We conducted a bootstrapping mediation analyses using the PROCESS macro (Model 4; Hayes, 2013) based on a resampling size of 10,000 with exceeded empathy expectations as the mediator and perceptions of support effectiveness as the dependent variable. The 95% bias-corrected confidence interval for the indirect effect of AI support through postsupport perceptions of exceeding empathy expectations on perceptions of support effectiveness (95% CI [.10, .54]) did not include zero, indicating that exceeded empathy expectations mediated the effect of support source (AI vs. human) on perceptions of support effectiveness (see Figure 7).

Alternative Mechanisms

We also explored two potential alternative mechanisms that could drive perceptions of support effectiveness. It is possible that one of the reasons people perceive AI support as more effective than human support is that they believe this form of support might help them deal with other issues in the future. Therefore, we began by exploring whether there were differences between conditions on this variable. An independent-samples *t* test revealed that participants in the AI support condition perceived the support giver as more likely to help deal with issues in the future ($M_{\text{AI Support}}$ = 3.34, *SD* = 1.23) compared with participants in the human support condition ($M_{\text{Human Support}} = 2.65$, $SD = 1.41$), $t(292) = 4.47$, $p < .001$, 95% CI of the difference [.39, .99], $d = .52$, 95% CI [.29, .76].

Figure 7

Exceeded Empathy Expectations Mediate the Effect of AI Support (vs. Human Support) on Perceptions of Support Effectiveness (Study 5)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
** $p < .01$. *** $p < .001$.

Similarly, it is possible that people feel that they are more able to express their negative emotions freely when disclosing to AI versus humans, and this could also lead them to perceive AI support as more effective than human support. We explored this possibility as well. An independent-samples *t* test revealed that participants in the AI support condition perceived that they could express their negative emotions more freely ($M_{\text{AI Support}}$ = 3.43, *SD* = 1.26) compared with those in the human support condition ($M_{\text{Human Support}}$ = 2.91, *SD* = 1.38), $t(292) = 3.36$, $p < .001$, 95% CI of the difference [.21, .82], $d = .39$, 95% CI [.16, .62].

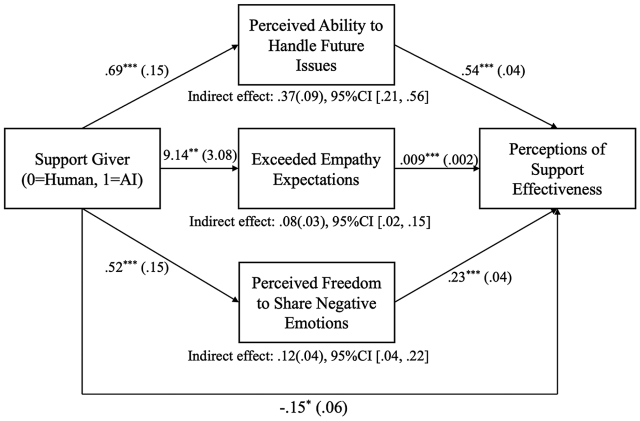
Mediation Analyses. To test whether exceeded empathy expectations remained a significant mediator when controlling for these alternative mechanisms, we conducted bootstrapping mediation analyses using the PROCESS macro (Model 4; Hayes, 2013) based on a resampling size of 10,000 with all three variables—exceeded empathy expectations, ability to help deal with future issues, and ability to share negative emotions freely—as parallel mediators, with perceptions of support effectiveness as the dependent variable. The indirect effect through exceeded empathy expectations remained significant (95% CI [.02, .15]), indicating that exceeded empathy expectations mediated the effect of support source on perceived effectiveness even when accounting for these alternative mechanisms. In addition, the indirect effects through ability to deal with future issues (95% CI [.21, .56]) and ability to share negative emotions freely (95% CI [.04, .22]) were also significant, suggesting that these variables may also contribute to perceptions of AI support effectiveness (see Figure 8).

Textual Analysis

We used the LIWC (Pennebaker et al., 2015) to explore whether participants' expression of negative emotions, anxiety, anger, and sadness differed across the AI and human conditions. We found no significant differences between the two conditions for negative emotions, $M_{\text{human}} = 4.47$, $SD_{\text{human}} = 2.92$; $M_{\text{AI}} = 4.81$, $SD_{\text{AI}} = 3.95$; $F(1, 292) = .68$, $p = .41$, $\eta^2 = .002$, 95% CI [.000, .026]; anxiety, $M_{\text{human}} = 1.13$, $SD_{\text{human}} = 1.54$; $M_{\text{AI}} = 1.18$, $SD_{\text{AI}} = 1.51$; $F(1, 292) = .07$, $p = .79$, $\eta^2 = .000$, 95% CI [.000, .014]; anger, M_{human} = 1.56,

Figure 8

Exceeded Empathy Expectations Mediate the Effect of AI Support (vs. Human Support) on Perceived Support Effectiveness (Even When Accounting for Two Additional Mechanisms; Study 5)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
* $p < .05$. ** $p < .01$. *** $p < .001$.

$SD_{human} = 1.69$; $M_{AI} = 2.06$, $SD_{AI} = 3.51$; $F(1, 292) = 2.33$, $p = .13$, $\eta^2 = .008$, 95% CI [.000, .040]; or sadness, $M_{human} = .42$, $SD_{human} = 0.79$; $M_{AI} = .44$, $SD_{AI} = 0.88$; $F(1, 292) = .05$, $p = .82$, $\eta^2 = .000$, 95% CI [.000, .014].

In sum, exceeded empathy expectations mediated the effect of AI support on perceived support effectiveness, and, importantly, this hypothesized indirect effect remained significant even when accounting for two alternative mechanisms. Participants also perceived AI as more helpful for resolving similar problems in the future and felt more able to express negative emotions freely when interacting with AI. These alternative mechanisms also partly drove the effect of AI support on perceived support effectiveness—a point we return to in the General Discussion section.

Study 6

Study 6 sought to replicate the empathy expectations effects observed in Studies 4 and 5 in a more ecologically valid setting. Participants interacted with either an actual human support giver or an AI chatbot in a live supportive exchange. We hypothesized that perceptions of AI (vs. human) support givers exceeding expectations for empathy would drive perceived support effectiveness.

Method

Participants

We recruited 401 employed U.S. adults (gender: 206 women, 194 men, one undisclosed; age: $M_{age} = 37.67$ years, $SD_{age} = 11.68$; ethnicity: 71 African American, 28 Asian/Asian American, 283 White, 28 Hispanic, three Arab/Arab American, seven Native American, three other) through Prolific Academic in November 2024. Participants were paid between $2.00 and $3.00 for the study (median completion time: approximately 10 min). This was a between-subjects experiment in which participants were randomly assigned to either a human support ($n = 200$) or AI support ($n = 201$) condition. A sensitivity analysis conducted using G*Power (Faul et al., 2007) indicated that with 401 participants, $\alpha = .05$ (two-tailed), and 80% power, this between-subjects design could detect effect sizes as small as $d = 0.28$. We preregistered this study on AsPredicted.org (see https://aspredicted.org/pxq7-rxt7.pdf).

Procedure

In this study, participants interacted in real time with their support giver, either a human or AI depending on the condition they were randomly assigned to.

AI Condition. To enable real-time interactions in the AI condition, we used the same customized AI chatbot that we developed and used in Studies 4 and 5, leveraging the GPT-4o model. Similar to previous studies, participants were informed that they would interact with an AI about a difficult social encounter and receive support. They wrote about the last time they had a negative encounter with another person and sent this as a chat message to the AI via our chat interface. Participants then received a supportive response from the AI and were able to converse back and forth. After this interaction, participants rated their perceptions of AI support.

Human Condition. To enable real-time interactions in the human condition, we again used ChatPlat. Unlike Study 3, where participants exchanged only a single message, participants in this study were able to converse back and forth with their human support giver. Participants were informed that they would interact with another person participating in this study about a difficult social encounter and receive support. They wrote about the last time they had a negative encounter with another person and sent this as a chat message to the human support giver via ChatPlat. Participants then received a supportive response and were able to converse back and forth with the support giver. After this interaction, participants rated their perceptions of human support.

Measures

Postsupport Perceptions About Empathy Expectations. Participants indicated the extent to which the support giver met their expectations for empathy using the same item as in Studies 4 and 5: "Please indicate the extent to which [the support giver] met your expectations for empathy": 0 (less empathetic than I expected), 50 (as empathetic as I expected), and 100 (more empathetic than I expected).

Perceptions of Support Effectiveness. We measured perceptions of support effectiveness using the same three-item scale used in Study 5, which assessed participants' perceptions of the extent to which they could count on the support to reduce their anxiety, stress, and negative affect after negative encounters ($\alpha = .96$).

Results

We examined whether people's perceptions that AI (vs. human) support exceeded their expectations for empathy mediated perceptions of support effectiveness. Thus, we expected and preregistered an indirect effect of support source on perceptions of support effectiveness through exceeded empathy expectations.

To examine this relationship, we conducted bootstrapping mediation analyses using the PROCESS macro (Model 4; Hayes, 2013) based on a resampling size of 10,000 with exceeded empathy expectations as the mediator and perceptions of support effectiveness as the dependent variable. Consistent with our prediction, results revealed that the 95% bias-corrected confidence interval for the indirect effect of AI support through exceeded empathy expectations on perceptions of support effectiveness (95% CI [.002, .31]) did not include zero, indicating that exceeded empathy expectations mediated the relationship between AI support and perceived effectiveness of this support (see Figure 9). In sum, in a real-time interactive support context, participants perceived that AI exceeded their expectations for empathy, which in turn drove perceptions of support effectiveness.[7]

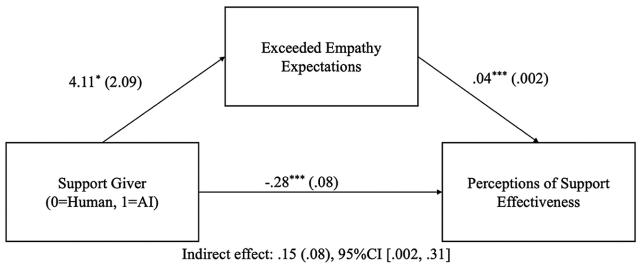
Textual Analysis

As in previous studies, we used LIWC (Pennebaker et al., 2015) to explore whether participants disclosed different types of information and expressed varying levels of distress to AI versus human support givers using dictionaries for negative emotion, anxiety, anger, and sadness. Results indicated that there were no significant differences between the two conditions in negative emotions, $M_{human} = 4.29$, $SD_{human} = 2.94$; $M_{AI} = 4.84$, $SD_{AI} = 3.14$; $F(1, 399) = 3.27$, $p = .07$, $\eta^2 = .008$, 95% CI [.000, .034]; anxiety, $M_{human} = 1.01$, $SD_{human} = 1.21$; $M_{AI} = 1.25$, $SD_{AI} = 1.68$; $F(1, 399) = 2.80$, $p = .10$, $\eta^2 = .007$, 95% CI [.000, .032]; anger, $M_{human} = 1.51$, $SD_{human} = 1.94$; $M_{AI} = 1.69$, $SD_{AI} = 2.00$; $F(1, 399) = .79$, $p = .38$, $\eta^2 = .002$, 95% CI [.000, .020]; or sadness, $M_{human} = .56$, $SD_{human} = 1.26$; $M_{AI} = .47$, $SD_{AI} = 1.01$; $F(1, 399) = .62$, $p = .43$, $\eta^2 = .002$, 95% CI [.000, .018], between the two conditions. Thus, our findings suggest that participants did not systematically differ what they shared with AI versus human support givers.

In sum, Study 6 provided converging evidence for the empathy expectations pathway in a real-time interactive support context. Participants who received support from an AI chatbot perceived that it exceeded their expectations for empathy, and this perception mediated the effect of AI support on perceived support effectiveness. These findings replicate the pattern observed in Studies 4 and 5, extending it to a more ecologically valid setting featuring live exchanges with actual human and AI support givers.

Figure 9

Exceeded Empathy Expectations Mediate the Effect of AI Support (vs. Human Support) on Perceptions of Support Effectiveness (Study 6)



Note. Values represent unstandardized coefficients, with standard errors in parentheses. CI = confidence interval.
* $p < .05$. *** $p < .001$.

General Discussion

Through six experiments (five preregistered), we examined how initial psychological expectations about AI versus humans shape people's experiences and perceptions of social support after minor aversive encounters. Studies 1–3 showed that participants experienced lower FNE when receiving support from AI (vs. humans), which in turn led to improved well-being outcomes. Studies 4–6 revealed that participants perceived AI as exceeding their initial expectations for empathy relative to human support givers, which drove perceptions of support effectiveness. These effects replicated across different support modalities (audio and text based), support giver characteristics (familiar and unfamiliar), and types of aversive encounters. Importantly, they emerged in both controlled simulation studies and more ecologically valid studies featuring interactions with AI and humans. Together, these findings indicate that initial psychological expectations about AI, particularly regarding its propensity for negative evaluation and capacity for empathy relative to humans, shape how people experience and perceive AI support.

Contributions to Theory

Our work makes several contributions to social psychological research. First, we contribute to the growing literature on the psychology of AI (Bigman et al., 2023; Newman et al., 2020; Raveendhran & Fast, 2021) and extend mind perception theory (H. M. Gray et al., 2007). Research drawing on mind perception theory has predominantly suggested that people respond more negatively to AI in social roles because AI is perceived as having lower experiential capacity—the ability to feel emotions and hold social motivations—than humans (Castelo et al., 2019; Epley & Waytz, 2010; Waytz et al., 2010). Our findings add nuance to this narrative. We show that in contexts wherein people seek support after minor aversive encounters, the very perception that AI lacks experiential capacity can be beneficial: Because people hold initial expectations that AI has a lower propensity for negative social evaluation, they experience lower fear of being judged, which in turn improves well-being. Rather than positioning AI's lower experiential capacity as uniformly detrimental in social contexts, our work reveals that when contexts activate concerns about interpersonal evaluation, perceived limitations in experiential capacity can become advantageous—a finding that extends the mind perception literature by specifying the conditions under which low experiential capacity benefits, rather than hinders, social interactions. Our work also extends emerging research demonstrating AI's functional effectiveness in support domains (De Freitas et al., 2025; Yin et al., 2024) by identifying the psychological mechanisms—specifically, expectations rooted in perceived experiential capacity—through which people's initial beliefs about AI shape whether support is experienced and perceived positively.

Second, we extend research on undersociality, which has largely focused on how people underestimate the positivity of human interactions, discouraging both support seeking and support giving (Dungan et al., 2022; Epley & Schroeder, 2014; Kardas et al., 2022; Kumar & Epley, 2023). We show that initial psychological expectations about AI—an entity increasingly capable of fulfilling

[7] There was no significant difference between AI and human support conditions on perceived support effectiveness ($p = .25$).

social functions—also shape support experiences but through a distinct mechanism. In human–human interactions, low expectations typically lead to missed opportunities or diminished outcomes. For AI, however, low initial expectations about its capacity for empathy (rooted in perceptions of AI's lower experiential capacity compared to humans) create a low bar that is more easily exceeded, enhancing how AI support is perceived. By demonstrating that lower expectations about nonhuman entities can foster unexpectedly positive support experiences, our research broadens the scope of undersociality research. Our findings highlight how human–AI interactions can sometimes overcome the very psychological barriers that limit human–human interactions in the social support context.

Finally, our work extends the social support literature by demonstrating that initial psychological expectations about the support giver, beyond the support itself, can shape its effectiveness. Prior research has focused on the presence, quality, or type of support (Mallinckrodt, 1989; Rini & Schetter, 2010; Taylor, 2011), emphasizing how features of the support influence outcomes. We show that initial expectations about the support giver's experiential capacity influence how support is experienced and perceived even when the support itself is comparable, highlighting that who (or what) delivers support matters independently of what is delivered.

Practical Implications

Our findings offer several practical implications. First, AI has the potential to provide effective episodic emotional support following the kind of minor, everyday aversive experiences examined here. While experiences like rude encounters with a supervisor or awkward interactions with friends may seem minor, they can still impair well-being, and people may hesitate to seek support from others out of fear of being judged or burdening them. AI can fill this gap by offering accessible, judgment-free support, though our findings speak specifically to brief, low-stakes interactions and do not address AI's role in more sustained or clinically significant support contexts wherein different dynamics may apply. This may be especially relevant for individuals who face barriers to accessing human support, whether due to cost, availability, or stigma, as AI can serve as a readily available supplement for brief emotional reassurance. Second, our findings suggest that people's initial expectations about AI's capacity for empathy are low relative to their expectations about humans. When AI exceeds these expectations, it enhances perceptions of support effectiveness. However, these initial expectations may shift over time as people gain more experience with AI—a possibility we discuss in the next section. Third, technologists should carefully consider people's psychological expectations about AI when designing support tools. Efforts to make AI overly humanlike may backfire by triggering concerns about judgment, undermining the very features that make AI feel psychologically safe in support contexts. Finally, organizations deploying AI in support roles—such as in education, wellness, or workplace settings—should recognize that AI may be particularly well-suited for contexts wherein people might otherwise avoid seeking support altogether. Our data do not suggest that AI should be limited to these contexts but rather that the psychological mechanisms we identified—reduced fear of evaluation and exceeded empathy expectations—are especially likely to operate when interpersonal barriers to human support are present.

Limitations and Future Directions

While our research offers consistent support for our theoretical model, our work is not without limitations (see Table 1) and opens several opportunities for future research. First, our theoretical framework rests on the assumption that people perceive AI as having lower experiential capacity than humans—that is, lower capacity to hold social motivations and experience emotions. This assumption likely holds for many current perceptions of AI, but it raises an important conceptual question: Are the benefits we observe driven by perceived capacity limitations or perceived motivational tendencies? Our studies do not disentangle this distinction. As AI systems become increasingly sophisticated, this distinction may become more consequential. For example, an AI that is perceived as technically capable of evaluating social behavior, but not motivated to do so, may produce different psychological dynamics than one perceived as fundamentally lacking such capacity or one that is fundamentally driven by social motivations to judge. Future research should directly measure and manipulate these distinct dimensions to understand how perceived capacity versus perceived motivation independently shape support experiences, particularly as public understanding of AI's capabilities evolves.

Second, and relatedly, our findings are grounded in people's initial expectations that AI has a lower propensity for negative social evaluation and a lower capacity for empathy compared to humans. These expectations may be relatively stable in the current moment, but they are unlikely to remain fixed. As people accumulate more experience with AI support tools, their expectations may calibrate upward, potentially changing how people experience and perceive AI support. Conversely, repeated positive interactions with human support givers might similarly reduce FNE if people come to see humans as less judgmental than initially expected. If expectations about both AI and human support givers shift with experience, the psychological gap between them, and the effects we observe, may narrow over time. Understanding how these initial expectations evolve with repeated interactions is a critical direction for future research, in particular, whether the benefits we observe with AI support are specific to initial encounters or persist as expectations recalibrate over time.

Third, while we identify exceeded empathy expectations as a key mechanism driving perceptions of AI support effectiveness, other factors may also play a role. Study 5 explored two alternatives—the belief that AI can better help deal with future aversive experiences and greater comfort expressing negative emotions to AI—and found that even when accounting for these, our primary mediator remained significant. However, these alternative mechanisms were also significant, pointing to promising directions for future research on additional psychological drivers of perceived support effectiveness.

Fourth, there are several methodological limitations worth noting. All participants were employed U.S. adults recruited via Prolific. It is possible that cultural factors, particularly norms around judgment, emotional expression, and support seeking, may moderate how psychological expectations shape support experiences. Future research should examine these dynamics across cultural contexts. Furthermore, in the two simulation studies (Studies 1 and 2), the support provider had a male voice, which may have influenced FNE dynamics; future work should systematically vary support giver characteristics, including gender. Study 3 also limited participants to a single message exchange, which may have constrained the support experience

Table 1
Summary of Limitations

Summary	Limitation	Solution
Capacity versus motivation	Our theoretical framework rests on the assumption that AI is perceived as having lower experiential capacity, but our studies do not disentangle whether benefits are driven by perceived capacity limitations or perceived motivational tendencies.	Future research should directly measure and manipulate perceived capacity versus perceived motivation, particularly as AI systems become more sophisticated and public perceptions evolve.
Initial expectations may shift.	Our findings are grounded in people's initial expectations about AI compared with humans. These expectations may calibrate over time with increased experience with both AI and human support.	Future research should examine how expectations about AI and human support givers evolve with repeated interactions and whether the benefits we observe are specific to initial encounters or persist over time.
Additional mechanisms	Although AI exceeding expectations for empathy was a key mediator in driving perceptions of support effectiveness, other alternative explanations may also contribute.	We addressed this limitation in Study 5 by exploring other possible mechanisms. Even when accounting for these possibilities, our primary mediator remained significant, but these findings point to promising directions for future work on additional psychological processes driving perceptions of support effectiveness.
Cultural generalizability	All participants were employed U.S. adults recruited via Prolific. Cultural factors, particularly norms around judgment, emotional expression, and support seeking, may moderate how psychological expectations shape support experiences.	Future research should examine these dynamics across cultural contexts.
Support giver characteristics	In the two simulation studies (Studies 1 and 2), the support provider had a male voice, which may have influenced fear of negative evaluation dynamics.	Future work should systematically vary support giver characteristics, including gender.
Conversational dynamics	Study 3 involved only a one-time message exchange between participants and support givers, which may have felt less natural, especially in the human condition.	While we addressed this limitation in Studies 4–6 by allowing back-and-forth conversation, future work should explore how conversational dynamics such as reciprocal exchange, responsiveness, depth of emotional disclosure, and interaction length shape support experiences with AI versus humans.
Self-report measures	All outcome measures were self-report. While appropriate for capturing subjective well-being and perceptions of support effectiveness, they do not speak to whether AI support influences actual behavior.	Future research should examine whether these effects extend to behavioral indicators such as greater willingness to seek support or subsequent coping strategies.
Support modality	Our studies exclusively examined support delivered through digitally mediated channels. While increasingly common, this does not capture face-to-face support dynamics wherein nonverbal cues and physical presence may activate different psychological processes.	We held modality constant across conditions to isolate the role of psychological expectations. Future research should examine how the modality of support delivery interacts with expectations about the support giver to shape support experiences.
Focus on episodic support and duration of effects	Our work focused on episodic emotional support following minor, everyday aversive experiences. It does not speak to the effectiveness of AI in delivering long-term or sustained support or support in response to more serious emotional challenges. Additionally, our studies measured outcomes immediately after the interaction, and we cannot speak to how long these effects persist.	Future research should examine whether the effects we observed operate differently when interactions extend beyond a single session and when people seek support repeatedly over time.
Potential moderators	Our work did not examine potential moderators of the effect of support source, such as the intensity, frequency, or timing of the aversive encounter.	Future work should consider factors that can shed light on when and for whom AI support can be most effective.

particularly in the human condition wherein a single exchange may feel less natural. Studies 4–6 addressed this by enabling back-and-forth conversation. Future research should explore how conversational dynamics such as the degree of reciprocal exchange and responsiveness of the support giver, depth of emotional disclosure, and interaction length shape support experiences with AI versus humans. Additionally, all outcome measures were self-report. While appropriate for capturing subjective experiences of well-being and perceptions of support effectiveness, self-report measures do not speak to whether AI support influences actual behavior such as greater willingness to seek support or subsequent coping strategies—an important direction for future research. Our studies also exclusively examined support delivered through digitally mediated channels—a context that, while increasingly common, does not capture face-to-face support dynamics wherein nonverbal cues and physical presence may activate different psychological processes. We held modality constant across conditions to isolate the role of psychological expectations, but future research should examine how the modality of support delivery impacts the support experience and the underlying mechanisms of the effects.

Fifth, our real-interaction studies used two distinct approaches to vary the support source. In Studies 3 and 6, participants' beliefs and actual interactions were aligned: those told they were interacting with a human did so and those told they were interacting with AI did so. In Studies 4 and 5, all participants interacted with an AI chatbot

but were led to believe they were interacting with either a human or AI. The consistency of our findings across both paradigms suggests that expectations about the support giver play a central role in shaping support experiences. However, we did not include explicit manipulation checks for whether participants believed they were interacting with a human or AI, as such checks could have introduced demand characteristics during the interaction. Future research could build on this using fully crossed designs—for example, having participants interact with a human while believing it is AI or vice versa—to isolate the independent contributions of beliefs about versus actual characteristics of the support giver. Developing unobtrusive methods for verifying participants' beliefs about their interaction partner would also be valuable.

Finally, our work focused on episodic emotional support following minor aversive experiences. It does not address the effectiveness of AI in providing sustained support or help with more serious emotional challenges, where the dynamics of support and its psychological effects may differ significantly. Additionally, our studies measured well-being and perceptions of support effectiveness immediately after the support interaction, and we cannot speak to how long these effects persist. Although our findings show that AI can offer effective episodic support, they should not be taken as evidence that AI can replace human support in more complex or ongoing contexts or that these effects are long lasting. Future research should examine whether the effects we observed operate differently when interactions extend beyond a single session and when people seek support repeatedly over time. It would also be valuable to examine how AI and human support might complement each other in various contexts and whether repeated brief interactions produce cumulative benefits.

Conclusion

AI is increasingly capable of providing social support, including after minor aversive experiences examined here. Our research demonstrates that how people experience and perceive AI support is shaped by the initial psychological expectations they bring to the interaction. Because people expect AI to have lower experiential capacity, they experience lower FNE, improving well-being, and perceive AI as exceeding their expectations for empathy, enhancing perceptions of support effectiveness. These findings underscore an important but overlooked insight: People's initial expectations of what AI can and cannot do fundamentally shape how they experience and perceive its support.

References

Adiwardana, D., Luong, M. T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G., Lu, Y., & Le, Q. V. (2020). *Towards a human-like open-domain chatbot*. ArXiv. https://arxiv.org/pdf/2001.09977

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Bigman, Y. E., Wilson, D., Arnestad, M. N., Waytz, A., & Gray, K. (2023). Algorithmic discrimination causes less moral outrage than human discrimination. *Journal of Experimental Psychology: General*, *152*(1), 4–27. https://doi.org/10.1037/xge0001250

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human–Computer Studies*, *59*(1–2), 119–155. https://doi.org/10.1016/s1071-5819(03)00018-1

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Crocker, J., & Canevello, A. (2008). Creating and undermining social support in communal relationships: The role of compassionate and self-image goals. *Journal of Personality and Social Psychology*, *95*(3), 555–575. https://doi.org/10.1037/0022-3514.95.3.555

Cutrona, C. E. (1990). Stress and social support: In search of optimal matching. *Journal of Social and Clinical Psychology*, *9*(1), 3–14. https://doi.org/10.1521/jscp.1990.9.1.3

Dakof, G. A., & Taylor, S. E. (1990). Victims' perceptions of social support: What is helpful from whom? *Journal of Personality and Social Psychology*, *58*(1), 80–89. https://doi.org/10.1037//0022-3514.58.1.80

Dang, J., & Liu, L. (2024). Extended artificial intelligence aversion: People deny humanness to artificial intelligence users. *Journal of Personality and Social Psychology*. Advance online publication. https://doi.org/10.1037/pspi0000480

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

De Freitas, J., Oğuz-Uğuralp, Z., Uğuralp, A. K., & Puntoni, S. (2025). AI companions reduce loneliness. *Journal of Consumer Research*, *52*(6), 1126–1148. https://doi.org/10.1093/jcr/ucaf040

Dungan, J. A., Munguia Gomez, D. M., & Epley, N. (2022). Too reluctant to reach out: Receiving social support is more positive than expressers expect. *Psychological Science*, *33*(8), 1300–1312. https://doi.org/10.1177/09567976221082942

Epley, N., Kardas, M., Zhao, X., Atir, S., & Schroeder, J. (2022). Undersociality: Miscalibrated social cognition can inhibit social connection. *Trends in Cognitive Sciences*, *26*(5), 406–418. https://doi.org/10.1016/j.tics.2022.02.007

Epley, N., & Schroeder, J. (2014). Mistakenly seeking solitude. *Journal of Experimental Psychology: General*, *143*(5), 1980–1999. https://doi.org/10.1037/a0037323

Epley, N., & Waytz, A. (2010). Mind perception. *Handbook of Social Psychology*, *1*(5), 498–541. https://doi.org/10.1002/9780470561119.socpsy001014

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, *4*(2), Article e7785. https://doi.org/10.2196/mental.7785

Gallup. (2024, June 12). *1 in 5 employees worldwide feel lonely*. https://www.gallup.com/workplace/645566/employees-worldwide-feel-lonely.aspx

Gallup. (2025). *World Happiness Report*. https://www.gallup.com/analytics/349487/world-happiness-report.aspx

Gleason, M. E. J., Iida, M., Shrout, P. E., & Bolger, N. (2008). Receiving support as a mixed blessing: Evidence for dual effects of support on psychological outcomes. *Journal of Personality and Social Psychology*, *94*(5), 824–838. https://doi.org/10.1037/0022-3514.94.5.824

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619. https://doi.org/10.1126/science.1134475

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, *125*(1), 125–130. https://doi.org/10.1016/j.cognition.2012.06.007

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hayes, A. F., & Preacher, K. J. (2013). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 219–266). Information Age Publishing.

Huang, K., Yeomans, M., Brooks, A. W., Minson, J., & Gino, F. (2017). It doesn't hurt to ask: Question-asking increases liking. *Journal of Personality and Social Psychology*, *113*(3), 430–452. https://doi.org/10.1037/pspi0000097

Jacobson, D. E. (1986). Types and timing of social support. *Journal of Health and Social Behavior*, *27*(3), 250–264. https://doi.org/10.2307/2136745

Jago, A. S., & Laurin, K. (2022). Assumptions about algorithms' capacity for discrimination. *Personality and Social Psychology Bulletin*, *48*(4), 582–595. https://doi.org/10.1177/01461672211016187

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, *33*(2), 125–143. https://doi.org/10.1177/0261927X13502654

Kanner, A. D., Coyne, J. C., Schaefer, C., & Lazarus, R. S. (1981). Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine*, *4*(1), 1–39. https://doi.org/10.1007/BF00844845

Karataş, M., & Cutright, K. M. (2023). Thinking about God increases acceptance of artificial intelligence in decision-making. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(33), Article e2218961120. https://doi.org/10.1073/pnas.2218961120

Kardas, M., Kumar, A., & Epley, N. (2022). Overly shallow? Miscalibrated expectations create a barrier to deeper conversation. *Journal of Personality and Social Psychology*, *122*(3), 367–398. https://doi.org/10.1037/pspa0000281

Kumar, A., & Epley, N. (2023). A little good goes an unexpectedly long way: Underestimating the positive impact of kindness on recipients. *Journal of Experimental Psychology: General*, *152*(1), 236–252. https://doi.org/10.1037/xge0001271

Leary, M. R. (1983). A brief version of the fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, *9*(3), 371–375. https://doi.org/10.1177/0146167283093007

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-u

Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, *37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Malle, B. F., Scheutz, M., Cusimano, C., Voiklis, J., Komatsu, T., Thapa, S., & Aladia, S. (2025). People's judgments of humans and robots in a classic moral dilemma. *Cognition*, *254*, Article 105958. https://doi.org/10.1016/j.cognition.2024.105958

Mallinckrodt, B. (1989). Social support and the effectiveness of group therapy. *Journal of Counseling Psychology*, *36*(2), 170–175. https://doi.org/10.1037/0022-0167.36.2.170

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, *31*(3), 301–306. https://doi.org/10.1111/j.2044-8260.1992.tb00997.x

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes*, *160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. University of Texas at Austin. https://doi.org/10.15781/T29G6Z

Porath, C., & Pearson, C. (2013). The price of incivility. *Harvard Business Review*, *91*(1–2), 114–121.

Qin, X., Zhou, X., Chen, C., Wu, D., Zhou, H., Dong, X., Cao, L., & Lu, J. G. (2025). AI aversion or appreciation? A capability–personalization framework and a meta-analytic review. *Psychological Bulletin*, *151*(5), 580–599. https://doi.org/10.1037/bul0000477

Raveendhran, R., & Fast, N. J. (2021). Humans judge, algorithms nudge: The psychology of behavior tracking acceptance. *Organizational Behavior and Human Decision Processes*, *164*, 11–26. https://doi.org/10.1016/j.obhdp.2021.01.001

Raveendhran, R., & Fast, N. J. (2024). When and why consumers prefer human-free behavior tracking products. *Marketing Letters*, *35*(3), 395–408. https://doi.org/10.1007/s11002-024-09726-6

Rini, C. M., & Schetter, C. D. (2010). The effectiveness of social support attempts in intimate relationships. In K. T. Sullivan & J. Davila (Eds.), *Support processes in intimate relationships* (pp. 26–68). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195380170.003.0002

Rizzo, A. A., Lucas, G., Gratch, J., Stratou, G., Morency, L. P., Shilling, R., & Scherer, S. (2016). Clinical interviewing by a virtual human agent with automatic behavior analysis. *2016 Proceedings of the international conference on disability, virtual reality and associated technologies* (pp. 57–64).

Rubin, M., Li, J. Z., Zimmerman, F., Ong, D. C., Goldenberg, A., & Perry, A. (2025). Comparing the value of perceived human versus AI-generated empathy. *Nature Human Behaviour*, *9*(11), 2345–2359. https://doi.org/10.1038/s41562-025-02247-w

Ruttan, R. L., McDonnell, M.-H., & Nordgren, L. F. (2015). Having "been there" doesn't mean I care: When prior experience reduces compassion for emotional distress. *Journal of Personality and Social Psychology*, *108*(4), 610–622. https://doi.org/10.1037/pspi0000012

Schilpzand, P., De Pater, I. E., & Erez, A. (2016). Workplace incivility: A review of the literature and agenda for future research. *Journal of Organizational Behavior*, *37*(Suppl. 1), S57–S88. https://doi.org/10.1002/job.1976

Schuetzler, R. M., Grimes, G. M., & Scott Giboney, J. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems*, *37*(3), 875–900. https://doi.org/10.1080/07421222.2020.1790204

Shockley, K. M., & Allen, T. D. (2013). Episodic work–family conflict, cardiovascular indicators, and social support: An experience sampling approach. *Journal of Occupational Health Psychology*, *18*(3), 262–275. https://doi.org/10.1037/a0033137

Spatola, N., & Urbanska, K. (2020). God-like robots: The semantic overlap between representation of divine and artificial entities. *AI & Society*, *35*(2), 329–341. https://doi.org/10.1007/s00146-019-00902-1

Stanton, J. M., Balzer, W. K., Smith, P. C., Parra, L. F., & Ironson, G. (2001). A general measure of work stress: The stress in general scale. *Educational and Psychological Measurement*, *61*(5), 866–888. https://doi.org/10.1177/00131640121971455

Taylor, S. E. (2011). Social support: A review. In H. S. Friedman (Ed.), *The Oxford handbook of health psychology* (pp. 189–214). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195342819.013.0009

U.S. Department of Health and Human Services. (2023). *Our epidemic of loneliness and isolation: The U.S. Surgeon General's advisory on the healing effects of social connection and community*. Office of the Surgeon General. https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf

Van Katwyk, P. T., Fox, S., Spector, P. E., & Kelloway, E. K. (2000). Using the Job-Related Affective Well-Being Scale (JAWS) to investigate affective responses to work stressors. *Journal of Occupational Health Psychology*, *5*(2), 219–230. https://doi.org/10.1037//1076-8998.5.2.219

Veiel, H. O. (1985). Dimensions of social support: A conceptual framework for research. *Social Psychiatry*, *20*(4), 156–162. https://doi.org/10.1007/BF00583293

Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the positive and negative affect schedule—Expanded form*. University of Iowa. https://doi.org/10.17077/48vt-m4t2

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Wolf, E. B., Lee, J. J., Sah, S., & Brooks, A. W. (2016). Managing perceptions of distress at work: Reframing emotion as passion. *Organizational Behavior and Human Decision Processes*, *137*, 1–12. https://doi.org/10.1016/j.obhdp.2016.07.003

Woolum, A., Foulk, T., & Erez, A. (2024). A review of the short-term implications of discrete, episodic incivility. *Social and Personality Psychology Compass*, *18*(1), Article e12918. https://doi.org/10.1111/spc3.12918

Yin, Y., Jia, N., & Wakslak, C. J. (2024). AI can help people feel heard, but an AI label diminishes this impact. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(14), Article e2319112121. https://doi.org/10.1073/pnas.2319112121

Received June 28, 2025
Revision received April 24, 2026
Accepted June 15, 2026 ■