CityU Scholars

Emotional Support from AI Chatbots
Should a Supportive Partner Self-Disclose or Not?

Meng, Jingbo; Dai, Yue Nancy

Publication record in CityU Scholars:
[Go to record](#)

Emotional Support from AI Chatbots: Should a Supportive Partner Self-Disclose or Not?

Jingbo Meng (ID) [1] & Yue (Nancy) Dai[2]

[1] Department of Communication, Michigan State University, East Lansing, MI 48824, USA

[2] Department of Media and Communication, City University of Hong Kong, Kowloon, Hong Kong

This study examined how and when a chatbot's emotional support was effective in reducing people's stress and worry. It compared emotional support from chatbot versus human partners in terms of its process and conditional effects on stress/worry reduction. In an online experiment, participants discussed a personal stressor with a chatbot or a human partner who provided none, or either one or both of emotional support and reciprocal self-disclosure. The results showed that emotional support from a conversational partner was mediated through perceived supportiveness of the partner to reduce stress and worry among participants, and the link from emotional support to perceived supportiveness was stronger for a human than for a chatbot. A conversational partner's reciprocal self-disclosure enhanced the positive effect of emotional support on worry reduction. However, when emotional support was absent, a solely self-disclosing chatbot reduced even less stress than a chatbot not providing any response to participants' stress.

Lay Summary

In recent years, AI chatbots have increasingly been used to provide empathy and support to people who are experiencing stressful times. This study compared emotional support from a chatbot compared to that of a human who provided support. We were interested in examining which approach could best effectively reduce people's worry and stress. When either a person or a chatbot was able to engage with a stressed individual and tell that individual about their own experiences, they were able to build rapport. We found that this type of reciprocal self-disclosure was effective in calming the worry of the individual. Interestingly, if a chatbot only reciprocally self-disclosed but offered no emotional support, the outcome was worse than if the chatbot did not respond to people at all. This work will help in the development of supportive chatbots by providing insights into when and what they should self-disclose.

Communicative artificial intelligence (AI) is an automated system that can perform communication tasks with some level of human intelligence (Frankish & Ramsey, 2014). One example of communicative AI is a chatbot that functions as an interpersonal interlocutor who converses with people via text-based communication. In recent years, chatbots (e.g., Woebot, Wysa) have been increasingly used to deliver mental health services and provide empathetic conversations to nonclinical populations (Miner, Milstein, & Hancock, 2017). This trend has created an urgency to deepen our understanding of using chatbots to provide emotional support to people during their stressful times.

When a person discloses his or her stressful experience and feelings, emotional support from a conversational partner can improve the person's psychological outcomes (Shenk & Fruzzetti, 2011). As AI chatbots begin to take part in empathetic conversations with people about their stress, important questions are raised as follows: Is a chatbot's emotional support effective at reducing people's stress and worry? When is its effectiveness amplified or diminished? According to the computers as social actors (CASA) framework (Nass & Moon, 2000), people should respond to a computer's emotional support in a way as if the support came from a human. Interpersonal communication theories claim that people's relational perceptions about a partner can enhance or impede their abilities to profit from the partner's emotional support (Weiss, 1980). Therefore, to achieve the effectiveness of emotional support, a chatbot should exhibit social cues that facilitate the formation of positive relational perceptions. Given that reciprocal self-disclosure enhances relational perceptions, such as liking and trust (Collins & Miller, 1994), this study translates reciprocal self-disclosure into a social cue of AI chatbots, which allow for the examination of how relational communication interacts with emotional support to influence people's psychological well-being. This study advances the CASA framework as applied in empathetic chatbots by articulating the boundary condition under which the effect of a chatbot's emotional support is more or less likely to occur.

If emotional support from AI chatbots helps to reduce people's stress and worry, does it mean that chatbots can replace humans for supportive communication during stressful times? Another aim of the present study is to examine how the source of support (i.e., chatbot vs. human) affects the effectiveness of emotional support and the moderating role of reciprocal self-disclosure in reducing stress and worry. Existing research has revealed that people present different social responses to humans and computers in the contexts of customer service (Sundar & Kim, 2019) and learning (Edwards, Edwards, Spence, Harris, & Gambino, 2016). The present study expands the comparison between humans and computers as conversational partners into the context of supportive communication and improving psychological well-being.

Literature review

Talking about stress: Emotional support

When a person discloses his or her stressful experience and feelings, for the person to gain psychological benefits, a conversational partner should support rather than ignore or blame the person (Shenk & Fruzzetti, 2011). Emotional support is a type of social support that communicates empathy, emotional validation, and encouragement to people who are experiencing stressful life events (Burleson, 2003). The provision of emotional support addresses basic human needs for being cared and supported by someone else. Emotional support has been shown to effectively reduce disclosers' stress and worry in face-to-face and computer-mediated interactions (Rains, Brunner, Akers, Pavlich, & Tsetsi, 2016). One pathway to explain the positive effect of emotional support is through perceived supportiveness of a partner, including whether a discloser feels that a partner's support is helpful and that the

partner can serve as a real source of support (Sarason, Sarason, & Shearin, 1986). When the perceived supportiveness is high, the discloser feels the partner's support is sensitive and thus, the partner can be a source of support when assistance is needed. These perceptions may encourage the discloser to reappraise a stressful situation as being less difficult because an effective source of support is available. As a result, the discloser's stress and worry are reduced, and psychological well-being is improved.

Although the above argument has been based on supportive communication via computers or face-to-face (Rains et al., 2016), the effect of emotional support and its pathway should manifest similarly when an AI chatbot serves as the support provider. According to the CASA framework (Reeves & Nass, 1996), people perceive and respond to computers as they do with humans in a natural way. People instinctively apply social scripts derived from experiences with humans to their interactions with computers. Recent studies have taken important steps toward having AI chatbots take part in the conversation with people about their stressful experiences (Ho, Hancock, Miner, 2018; van der Zwaan, Dignum, Jonker, & van der Hof, 2014). For example, people experienced significant emotional benefits (i.e., feeling better) after receiving emotional support from a chatbot (Ho et al., 2018). Similarly, a virtual buddy named Robin could express sympathy to the victims of cyberbullying. The victims considered Robin as a caring supporter and reported reduced distress (van der Zwaan et al.). Therefore, we hypothesize that the effect of emotional support and its pathway onto a discloser's psychological well-being will manifest no matter if the conversational partner is a human or a chatbot.

> H1: A discloser reduces more (a) stress and (b) worry when a conversational partner provides emotional support than when the partner does not.
> H2: The positive effect of emotional support on (a) stress and (b) worry reduction is mediated through perceived supportiveness of the partner.

Chatbot versus human as a conversational partner

The source of messages is one of the most enduring subjects in human communication research. Individuals actively orient themselves toward the source of messages, which may affect psychological outcomes after receiving the messages. Although the CASA framework argues that people respond to computers as if they were social actors (Reeves & Nass, 1996), it does not claim that people would treat computers exactly the same as real humans in every setting. Existing research has revealed that people show different responses to computers versus humans in the contexts of learning (Edwards et al., 2016) and completing a service task (Sundar & Kim, 2019). What drives different responses is the machine heuristic that refers to mental shortcuts wherein people attribute machine characteristics when making judgments about an interaction (Sundar, 2008). When the source of interaction is a machine (e.g., computer), people automatically apply stereotypes about a computer such that it is mechanistic, objective, unemotional, and cold (Sundar & Kim, 2019), which in turn, shape the outcome of interactions.

In the context of stress talks, a discloser needs to feel that the conversational partner truly understands and cares his or her situation before psychological gains could occur (Reis, Lemay, & Finkenauer, 2017). Compared with human partners, chatbots may trigger machine heuristic during initial interactions. People apply the mental shortcut such that chatbots are not able to feel human emotions and their responses are programmed. The algorithm-based emotional support may appear to be scripted and ungenuine. People are then less likely to consider chatbots as real sources of emotional support. In contrast, human partners are believed to have the ability of empathy and their emotional support reflects true understanding and caring (Stein & Ohler, 2017). Therefore, human

partners are more likely to be considered as real sources of support and receiving emotional support from a human partner may be more beneficial than from a chatbot.

H3: Disclosers perceive higher supportiveness from a human partner than a chatbot who provides emotional support, which further leads to (a) stress and (b) worry reduction.

Boundary condition: Reciprocal self-disclosure

A natural question to ask next is, under which condition, an AI chatbot's emotional support could more effectively improve people's psychological well-being. According to the CASA framework, social cues can invoke the schemata of human–human interaction. The presence of certain social cues may enhance the effectiveness of a chatbot's emotional support during stress talks. Social cues refer to a computer's talks (via either text or speech) that convey the computer's personality and that apply social dynamics (e.g., reciprocity) as in human–human interactions (Nass & Moon, 2000). Reciprocal self-disclosure can be a social cue for a chatbot to simulate human–human conversations, as reciprocity is one of the most frequently observed norms during interpersonal communication (Gouldner, 1960).

Reciprocal self-disclosure is an act that a conversational partner reveals personal information equally intimate as what the discloser reveals (Hill & Stull, 1982). When a partner discloses reciprocally, the equity in the relationship is maintained between the discloser and the partner, and thus, the discloser finds the partner more likable and trustworthy (Collin & Miller, 1994). Research on interpersonal communication has long recognized that although receiving emotional support can yield psychological benefits, personal relationships may amplify or impede people's ability to profit from their partners' emotional support (Jones & Burleson, 1997). The prevailing relational perception, such as equity and trust in the relationship, exerts an influence on the effect of the partner's supportive behavior (Weiss, 1980). Evidence from therapist–client communication has shown that clients gave higher ratings regarding likeability, sincerity, and strength of the relationship to therapists who disclose about themselves, which enhance the success of the therapists' support (McCormic et al., 2014).

According to the CASA framework, reciprocal self-disclosure as a social cue renders the human–human social script more accessible and applicable. Like in interpersonal communication, a chatbot's self-disclosure may affect people's relational perceptions about the bot. People may perceive a self-disclosing chatbot as more sincere and trustworthy. This perception could create a relational environment that is more conducive to the positive effect of receiving emotional support from the chatbot on psychological well-being. In other words, emotional support from a chatbot should have a stronger positive effect on reducing stress and worry when the chatbot discloses to establish a trusting relationship. On the basis of above arguments, we propose a specific moderating effect of reciprocal self-disclosure below.

H4: A conversational partner's reciprocal self-disclosure will magnify the positive effect of the partner's emotional support on reducing a discloser's (a) stress and (b) worry.

When comparing humans and chatbots as conversational partners, we speculate a different magnitude of the moderating effect of reciprocal self-disclosure. During initial interactions, knowing the source of communication affects how people engage in relational communication and ultimately benefit from a partner's emotional support. Past work has documented that compared with those who interact with computers, people who interact with human partners use more relationship-related statements, such as self-disclosure (Mou & Xu, 2017), reference to the partner (e.g., thank you), and

taking advice from the partner (e.g., I should have. . .) (Shechtman & Horowitz, 2003). These findings suggest that relational communication comes more naturally when the partner is a human than a computer. Machine heuristic may have led people less likely to consider relational attributes of a chatbot during initial interactions, so a chatbot has a greater need than a human to proactively engage in relational communication via reciprocal disclosure to compensate for the stereotypes about machines. Likewise, according to the CASA framework, the effect of a chatbot's emotional support is a function of the extent to which a social script for human–human interaction is triggered and followed. A chatbot should be more dependent on the presence of the social cue (i.e., reciprocal self-disclosure) whereas a human supporter does not need this social cue to convey the meanings of their utterances. Therefore, reciprocal self-disclosure from a human may not be as important as it is for a chatbot to enhance the effectiveness of emotional support.

> H5: The magnifying effect of reciprocal self-disclosure on the relationship between emotional support and reduction in (a) stress and (b) worry should be stronger when the source is a chatbot than a human.

Method

Overview

The study featured a web-based experiment with a 2 (source: chatbot vs. human) by 2 (emotional support: yes vs. no) by 2 (reciprocal self-disclosure: yes vs. no) between-subjects factorial design. Participants were randomly assigned to one of eight conditions. Our chatbots were built using an existing AI developing tool called Chatfuel (www.chatfuel.com), which can recognize keywords from user inputs and respond using those keywords.

Procedure

Upon starting the questionnaire, participants were prompted to think of an issue that has been stressful to them lately and their perceived stress and worry about the issue were assessed. They were then instructed that they would have a conversation with either an AI chatbot or a person about the stressful situation in mind. Embedded in the online questionnaire, Chatfuel activated a participant's Facebook messenger as a pop-up window to start the chat. The chatbot and human partners followed the same predefined script (Appendix A in Supporting Information) with slight differences in reciprocal disclosure messages so that the disclosure reasonably fit the identity of a chatbot or a human partner. After the chat, participants completed the questions about their perceived support effectiveness of the partner, perceived stress and worry regarding the stressful issue, and manipulation checks.

The chat started with the conversational partner (i.e., chatbot or human) initiating a question to elicit self-disclosure from a participant about his/her stressful issue. After the participant disclosed, depending on the experimental condition, the chat partner provided no response (i.e., control) or provided feedback with emotional support and/or reciprocal disclosure. This completed one turn for the chat. There were in total six turns for the entire chat. To enhance the manipulation of conversational partners, we have adopted the presence of "dot dot dot" suggesting the partner was typing is human conditions. In addition, a human partner's response time was 5-second longer than a chatbot's, assuming that the chatbot was able to provide faster replies than the human partners typing words.

Participants

Participants were recruited from a subject pool at a large Midwestern university. A total of 278 participants took part in the study. Participants were excluded if they did not complete survey questions ($n = 29$), and if they did not have a chat record ($n = 26$, possible reasons: technical issues that failed logging their chats or participants may have skipped the chat and continued with the survey). After reading through their chat records, participants were also excluded if they failed to follow directions such as missing a turn in the conversation ($n = 4$) or responded using emoticons instead of narratives ($n = 3$). Moreover, participants were excluded when they attempted to test the bot by asking irrelevant questions ($n = 5$). The final sample consisted of 211 cases. Participants' average age was 20.4 ($SD = 2.28$) and 61.6% were females.

Measurement

Perceived stress

Perceived stress scale measures the degree to which situations in one's life are perceived as stressful in terms of negative affective reactions and lack of control (Cohen, Kamarck, & Mermelstein, 1983). Participants were asked to rate their feelings about the stressful issue self-identified. Example items included "I feel upset," "I feel nervous and stressed," and "I feel that difficulties are piling up so high that I can't overcome them." Participants rated the statements on a 7-point scale ($1 = strongly\ disagree$, $7 = strongly\ agree$), Cronbach's $\alpha_{pre} = .77$, Cronbach's $\alpha_{post} = .80$. Stress reduction was computed by subtracting participants' post-test stress scores from their pretest scores, $M_{pre} = 4.25$, $SD_{pre} = .81$, $M_{post} = 3.90$, $SD_{post} = .87$.

Worry

Participants were asked how much worry they had about the stressful situation they have identified. This measure was adapted from Rains et al. (2016). Participants answered the question on a 1–7 Likert scale ($1 = not\ at\ all\ worried$, $7 = extremely\ worried$). Reduction of worry was computed by subtracting participants' post-test worry scores from their pretest worry scores, $M_{pre} = 4.94$, $SD_{pre} = 1.20$, $M_{post} = 4.54$, $SD_{post} = 1.38$.

Perceived supportiveness of a partner

Nine items adapted from Zimet, Dahlem, Zimet, and Farley (1988) measured participants perceived supportiveness on a 7-point scale ($1 = strongly\ disagree$, $7 = strongly\ agree$), Cronbach's $\alpha = .90$, $M = 3.94$, $SD = 1.25$. Example items from the scale included: "I can get the emotional help and support I need from my chat partner," and "My chat partner was a real source of comfort to me."

Control variables

Neuroticism is one of the Big Five personality traits that describe individuals' tendency toward negative feelings. Neuroticism was included as a covariate because individuals with high-neuroticism experience greater distress and worry in response to stressful life events (Gunthert, Cohen, & Armeli, 1999). Neuroticism was measured on a 7-point scale ($1 = strongly\ disagree$, $7 = strongly\ agree$) (Rammstedt & John, 2007). Example items include, "I see myself as someone who gets nervous easily." $M = 4.31$, $SD = 1.25$, Cronbach's $\alpha = .76$.

Manipulation check questions

To assess the manipulations of emotional support and reciprocal disclosure, participants were asked whether ($1 =$ yes, $2 =$ no) their conversational partner comforted them and reciprocally disclosed his/her stressful issue in two separate questions. In addition, we examined participants' language style to test whether they really believed they were conversing with a chatbot or a human as instructed (Ho et al., 2018). Previous research has identified that people used more netspeak words (e.g., btw, lol, thx) and informal language (e.g., fillers, nonfluencies such as "hmm" or "umm") in human–human online conversations than in human–chatbot conversations (Hill, Randolph Ford, & Farreras, 2015; Ho et al., 2018). In addition, people tend to use shorter but more sentences when interacting with computers than with humans (Branigan, Pickering, Pearson, & McLean, 2010). Shorter sentences have also been found to contain fewer conjugations (e.g., "and," "whereas"; Tausczik & Pennebaker, 2010). Therefore, we measured netspeak words, informal language, and conjugations through LIWC.

Results

Manipulation check

Two chi-squared tests assessed the manipulations of emotional support and reciprocal disclosure, respectively. The induction of emotional support was significantly associated with participants' reports of whether their conversational partners comforted them during the chat, χ^2 $(1, N = 183) = 71.72$, $p < .001$, Cramer's $V = .63$. The induction of reciprocal disclosure was also significantly associated with participants' reports of whether their conversational partners self-disclosed during the chat, χ^2 $(1, N = 183) = 127.93$, $p < .001$, Cramer's $V = .84$.

Three *t*-tests compared the use of netspeak words, informal language, and conjugations between participants assigned to the human vs. chatbot conditions. The analyses revealed that participants used more netspeak ($M_{human} = .26$, $SD_{human} = 1.03$; $M_{bot} = .02$, $SD_{bot} = .14$; $t = -2.48$, $p < .01$), more informal language ($M_{human} = .96$, $SD_{human} = 2.47$; $M_{bot} = .48$, $SD_{bot} = 1.0$; $t = -1.99$, $p = .02$), and more conjugations ($M_{human} = 1.12$, $SD_{human} = .99$; $M_{bot} = .05$, $SD_{bot} = .63$; $t = -2.51$, $p < .01$) when chatting with a human partner than with a chatbot. These statistics suggested that during the actual interaction, participants acted in the way as expected when interacting with a human versus a chatbot.

Hypothesis testing

To ensure equivalence among different experimental conditions, we have compared group differences on demographic variables and neuroticism that may affect perceived stress and worry (Gunthert et al., 1999). Although conditions did not differ on demographic variables, participants in the human conditions had a higher level of neuroticism ($M = 4.52$, $SD = 1.26$) than those in the chatbot condition ($M = 4.12$, $SD = 1.21$, $t = -2.28$, $p = .03$). We hence included neuroticism as a covariate in the analyses.

H1 posited a main effect of emotional support on (a) stress and (b) worry reduction. An analysis of covariance (ANCOVA) analysis showed that emotional support did not have a significant effect on stress reduction, $F(1, 169) = 1.06$, $p = .31$, $\eta_{p}^2 = .01$, but had a positive effect on worry reduction, $F(1, 169) = 6.84$, $p = .01$, $\eta_{p}^2 = .04$. Participants reduced more worry when receiving emotional support ($M = .48$, $SD = .96$) than not ($M = 0.12$, $SD = 1.02$, $p = .01$). Neuroticism also significantly predicted worry reduction, $F(1, 169) = 5.00$, $p = .03$, $\eta_{p}^2 = .03$.

H2 posited a mediation relationship between emotional support and stress/worry reduction through perceived supportiveness. Hayes' PROCESS Macro (Hayes, 2013) was used to assess this mediation effect. The results showed a significant *indirect* effect of receiving emotional support on stress reduction through perceived supportiveness of, $b = .12$, $SE = .05$, bias-corrected 95% confidence interval (CI) [.04, .23]. Specifically, receiving emotional support increased a discloser's perceived supportiveness ($b = .99$, $t = 5.67$, $p < .001$), which in turn, predicted stress reduction ($b = .13$, $t = 2.97$, $p = .003$). The *direct* effect of receiving emotional support on stress reduction was not significant, $b = -.02$, $SE = .11$, $p = .83$, bias-corrected 95% CI [–.23, .19]. The total effect of emotional support on stress reduction was not significant, $b = .10$, $SE = .10$, $p = .31$. Taken together, these results indicated emotional support exerted a positive influence on stress reduction only indirectly through perceived supportiveness. Therefore, H2a was supported.

The mediation analysis also showed a significant *indirect* effect of emotional support on worry reduction through perceived supportiveness, $b = .21$, $SE = .08$, bias-corrected 95% CI [.07, .37]. Specifically, receiving emotional support increased a discloser's perceived supportiveness ($b = .97$, $t = 5.56$, $p < .001$), which further predicted worry reduction ($b = .21$, $t = 3.38$, $p < .001$). The *direct* effect of emotional support on worry reduction was not significant, $p = .38$, bias-corrected 95% CI [–.40, .38]. The total effect of emotional support on worry reduction was significant, $b = .34$, $SE = .15$, $t = 2.31$, $p = .02$, bias-corrected 95% CI [.05, .63]. Therefore, H2b was supported.

H3 posited a moderating role of the source in the mediation relationship. Hayes' PROCESS Macro (Model 7, Hayes, 2013) was used to test the hypotheses. The moderated mediation index suggested that the conditional indirect effects of emotional support on stress reduction was different depending on the source (Index $= .08$, $SE = .05$, 95% CI [0.003, .19]). Specifically, a human partner's emotional support had a relatively stronger impact on perceived supportiveness of the partner ($b = 1.32$, $SE = .25$, $p < .001$, 95% CI = [0.83, 1.81]) than a chatbot's ($b = 0.67$, $SE = .24$, $p = .01$, 95% CI = [0.18, 1.15]). Similarly, the conditional indirect effects of emotional support on worry reduction were different depending on the source (index $= .17$, $SE = .10$, 95% CI [.02, .39]). Specifically, a human partner's emotional support had a relatively stronger impact on perceived supportiveness ($b = 1.37$, $SE = .25$, $p < .001$, 95% CI = [.88, 1.86]) than a chatbot's ($b = .58$, $SE = .24$, $p = .02$, 95% CI = [.10, 1.06]). Figure 1 presents specific coefficients for the moderated mediation models.

H4 and H5 were tested with two ANCOVA models using stress and worry reduction as the dependent variables, respectively, the experimental conditions and interactions as the predicting factors, and neuroticism as the control variable. With respect to the two-way interaction in H4, for stress reduction, the results did not reveal any significant interaction between reciprocal disclosure and emotional support, $F(1, 163) = 2.08$, $p = .15$, $\eta_p^2 = .01$. H4a was not supported. For worry reduction, the analysis revealed a significant interaction between reciprocal disclosure and emotional support, $F(1, 163) = 5.48$, $p = .02$, $\eta_p^2 = .03$. Specifically, participants experienced greater worry reduction when reciprocal disclosure was provided in addition to emotional support ($M = 0.64$, $SD = 1.08$) than those in emotional support only conditions ($M = .30$, $SD = .83$, $p = .04$, one-tailed[1]). These results supported H4b (Figure 2).

With respect to the three-way interaction in H5, for stress reduction (H5a), although the three-way interaction was not statistically significant, $F(1, 169) = 3.41$, $p = .07$, $\eta_p^2 = .02$, its p-value was close enough to the threshold of significance to justify subsequent post hoc analyses to further understand the nature of the interaction. The analyses showed a combined effect between reciprocal disclosure and emotional support on stress reduction for a chatbot but not for a human partner. However, the results revealed a significant interaction pattern different from what was proposed in H5a.
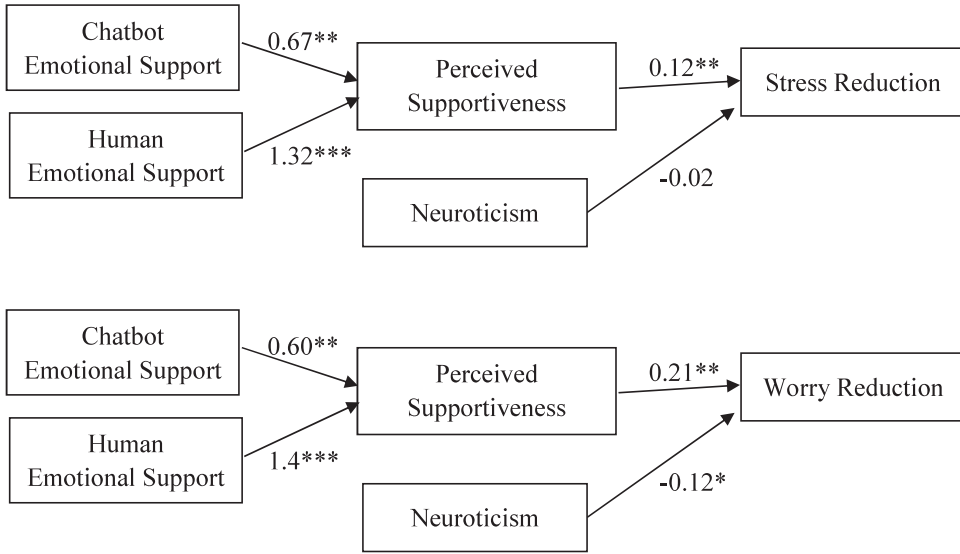
Note. *** *p* < .001, ** *p* < .01, * *p* < .05. Coefficients are unstandardized

Figure 1 Path coefficients from moderated mediation analyses on stress and worry reduction. ***p < .001, **p < .01, *p < .05. Coefficients are unstandardized.
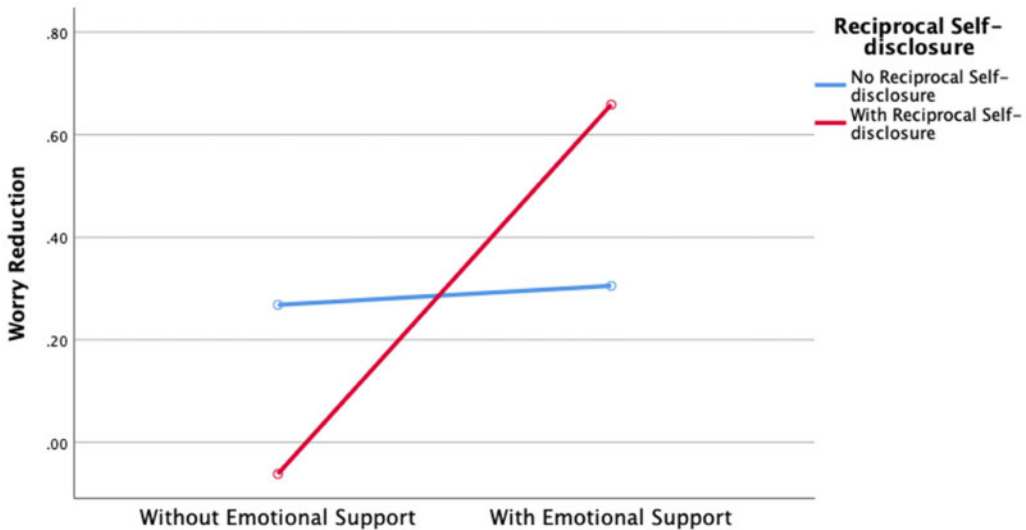
Figure 2 Interaction effect between emotional support and reciprocal disclosure on worry reduction. The effects are estimated at the average value of the covariate (neuroticism = 4.31).

Participants did not experience more stress reduction when reciprocal disclosure was provided in addition to emotional support ($M = .51$, $SD = .73$) than those in emotional support only conditions ($M = .33$, $SD = .53$, $p = .35$). Instead, when a chatbot did not provide emotional support, not

disclosing in reciprocation ($M = .50$, $SD = .52$) resulted in more stress reduction than disclosing ($M = .05$, $SD = .72$, $p = .02$). In contrast, for human partners, we did not observe the same interaction effect. When a human partner provided emotional support, disclosing in reciprocation ($M = .38$, $SD = .70$) did not result in more stress reduction than not disclosing ($M = .34$, $SD = .68$, $p = .75$); when the human partner did not provide emotional support, there was also no difference in stress reduction whether they self-disclosed ($M = .37$, $SD = .59$) or not ($M = .21$, $SD = .75$, $p = .47$). These results suggested that a chatbot that only self-disclosed without providing emotional support was the worst at reducing stress, but offering emotional support could offset the negative effect of its reciprocal self-disclosure (Figure 3). For worry reduction (H5b), there was not any significant three-way interaction, $F(1, 169) = 1.47$, $p = .23$, $\eta_{\mathrm{p}}^2 = .01$, meaning that the combined effect of emotional support and reciprocal disclosure did not differ between human and chatbot partners.

Discussion

The present study examined how and when chatbots' emotional support was effective in reducing people's stress and worry during conversations about their stressful experiences. The results explained the process and the boundary condition under which empathetic chatbots could help to improve people's psychological well-being. We have found that emotional support from both a chatbot and a human partner contributed to stress and worry reduction fully mediated through perceived supportiveness of the partner. The finding is consistent with the literature on supportive communication in dyadic interactions (Priem & Solomon, 2015), highlighting that a successful chatbot for stress talk needs to make people believe that the chatbot could serve as a real and reliable source of support. This could be achieved by designing high quality of support messages or other social cues not yet tested in the present study (Rains et al., 2019).

When comparing human and chatbot as a source of emotional support, we found that emotional support from a human partner led to greater perceived supportiveness of the partner than that from a chatbot. The finding resonates with the argument about machine heuristic (Sundar, 2008), such that people apply stereotypes about a chatbot when interpreting their interactions. Compared with human partners, chatbots are less capable of feeling and relating (Edwards et al., 2016). Although this study did not explicitly frame the conversation as supportive communication, participants were prompted to disclose their personal stress first. Therefore, the stressor was a clear antecedent to the conversation, which may have shaped participants' judgments about whether or not chatbots could serve as real sources of support. Compared with chatbots, the same emotional support messages coming from human partners may be perceived as more sensitive and genuine, and thus, human partners were considered as more helpful sources of support to reduce stress and worry.

With respect to reciprocal self-disclosure as a boundary condition for the effect of emotional support, we have found different patterns for worry and stress reduction. For worry reduction, a partner's emotional support resulted in a greater worry reduction when the partner disclosed about his- or herself than when the partner who did not. Worry is a chain of negative thoughts surrounding a stressor or a personal problem (Borkovec, Alcaine, & Behar, 2004). Worry seeks answers to questions such as "What is the worst thing that could happen?" and "Will others judge me or feel disappointed in me?" (Capobianco, Morris, & Wells, 2018). By self-disclosing his or her own stressful experience, a conversational partner showed vulnerability and reciprocity. This act may have made participants feel a more fair and transparent relationship. Then, the partner's comforting words could more effectively calm down participants' negative thoughts. An alternative explanation is that in a more fair and
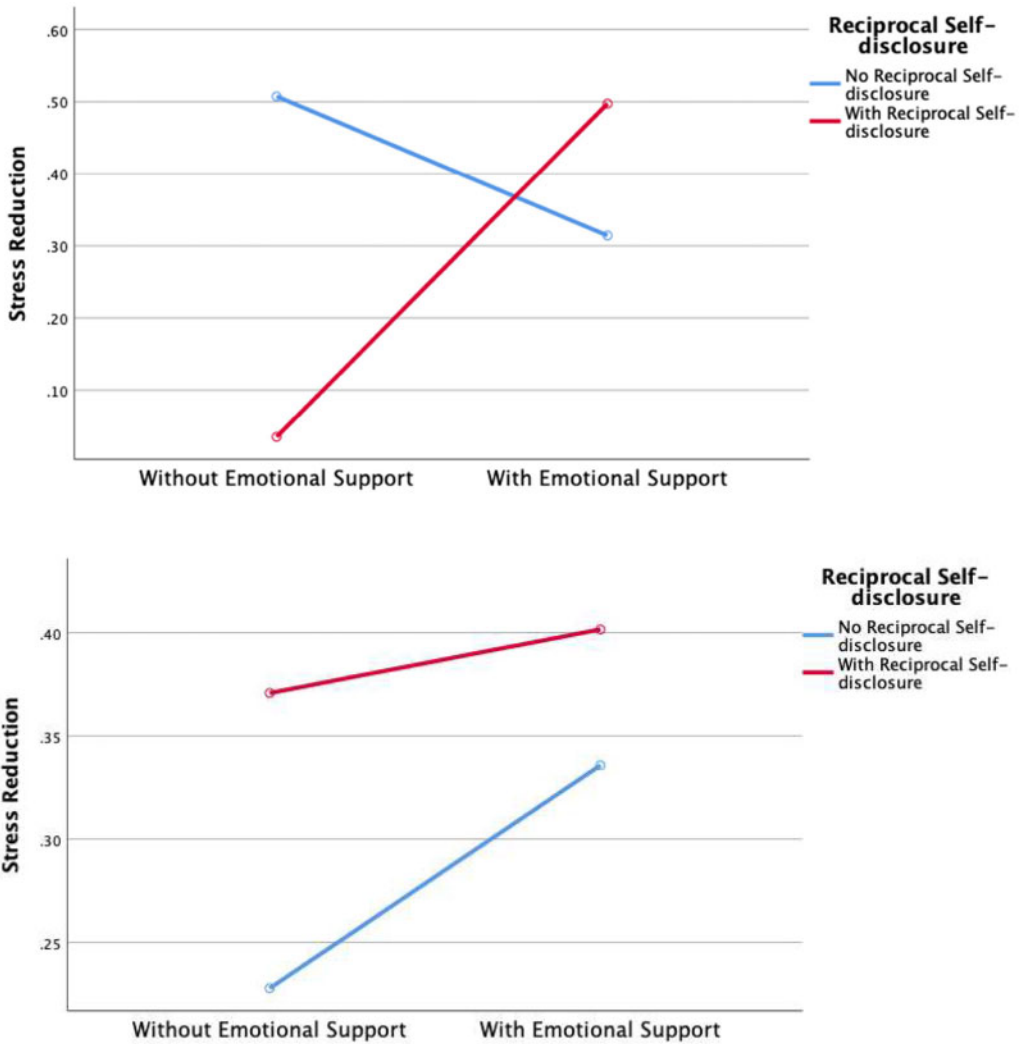
Figure 3 Interaction effect between emotional support and reciprocal disclosure on stress reduction after interacting with a chatbot (top) and with a human (bottom). The effects are estimated at the average value of the covariate (neuroticism = 4.31).

transparent relationship, participants could engage in or enjoy the conversation more. The enjoyment of the interaction may have simply distracted participants from their concerns about the stressor. Thus, they were less self-focused and had fewer thoughts about the stressor.

In contrast, we did not find the same interaction effect between reciprocal self-disclosure and emotional support for stress reduction. Perceived stress is the degree to which situations in one's life are appraised as stressful, overloading, and uncontrollable (Cohen et al., 1983). The experienced level of stress is a function of objective stressful events and subjectively appraised coping resources. It is possible that although participants could suspend their thoughts about the stressful issue for the

moment, a conversational partner's emotional support was not powerful enough to change the occurrence of the stressful event or significantly enhance participants' coping resources. Therefore, the partner's emotional support alone, or combined with reciprocal disclosure, did not have a *direct* effect on reducing perceived stress.

For stress reduction, the findings, instead, showed a distinct interaction effect from our hypotheses. A chatbot's reciprocal self-disclosure alone had a negative effect on stress reduction such that it reduced even less stress than the control condition wherein the chatbot did not provide any response to participants' disclosure about their stress. But this negative effect was significantly mitigated by the chatbot's emotional support. In contrast, a human partner's reciprocal self-disclosure alone did not have a negative effect on stress reduction. The comparison between chatbot and human suggested that the source of support may have altered the meaning of reciprocal self-disclosure. When stressful experience was a clear antecedent to the conversation, participants may expect to receive emotional comfort from their partners. Without emotional support provided, a chatbot's reciprocal self-disclosure may sound irrelevant and surreal to participants (i.e., stress due to being an incapable bot). A solely self-disclosing chatbot may make participants feel their stressful feelings were not attended to at all. Therefore, a chatbot's reciprocal self-disclosure alone had a backfire effect on stress reduction. However, a human partner's reciprocal self-disclosure may sound more relatable because of common human experience (e.g., stress due to looking for a job). Indeed, a human partner's disclosure about similar feelings could be interpreted as a form of showing understanding (Burleson, 2003). Although this form of understanding was not strong enough to significantly reduce stress, it did not backfire as observed in the chatbot condition.

Further theoretical reflections

Our findings supported the CASA framework by showing similar effects of a human's and a chatbot's feedbacks on perceived supportiveness and worry reduction. However, we found a different pattern of interaction between emotional support and reciprocal disclosure on stress reduction when comparing a chatbot and a human partner. These findings direct our attention to a theoretical concern about different levels of social responses triggered by social cues. A chatbot's emotional support is able to activate people's positive perceptions (e.g., caring, supportiveness) about the bot (Liu & Sundar, 2018; van der Zwaan et al., 2014), but its direct effects on psychological benefits are subject to variations (Ho et al., 2018). As this study reveals, emotional support and reciprocal disclosure are effective in calming down worrisome thoughts, but not for *directly* reducing perceived stress. Suspending worries could be a more immediate or temporal benefit from a pleasant conversation but reducing perceived stress requires a complex appraisal of the stressful situation and one's coping resources (Rains et al., 2019). More investigations are needed to understand how social cues affect perceptual responses, and immediate and distal psychological responses differently.

Another theoretical concern is related to the context-dependent meaning of social cues, emphasizing a match between the context and the choice of social cues. In the context of stress talks, emotional support is a critical baseline behavior to offer. When emotional support is missing, people may engage in a motivated interpretation of the social cue (i.e., reciprocal self-disclosure) based on their needs for empathy. A chatbot's surreal self-disclosure could still contribute to an enjoyable conversation, but it reinforces machine heuristic that chatbots lack the ability of empathy (Sundar, 2008; Reis et al., 2017). Therefore, a chatbot's reciprocal disclosure alone may not have triggered human-human schemata for stress talks. It is worth noting that the negative effect of a chatbot's self-disclosure on stress reduction does not necessarily

contradict with CASA but rather emphasizes designing the chatbot's disclosure messages more equivalent to humans' disclosure. In other words, social cues should be designed in a way to trigger specific aspects of human–human schemata that are contextually meaningful (Rains et al., 2019).

Limitations and future research

The first limitation in this study is that emotional support messages provided by humans and chatbots were relatively static and could not fully reflect natural conversations. Humans and chatbots may talk the same content with varying linguistic features such as word choices and sentence structures. Future studies should consider those linguistic features when manipulating the sources of support. Moreover, to enhance the manipulation of support providers, the response time was 5-second longer for human partners than for chatbots. Immediate responses may indicate a partner's warmth and accessibility but may also imply the partner's responses are not well-thought or wholehearted. Future research is encouraged to examine expected response time and how it interacts with the source to affect their perceived support and other outcomes.

The second limitation involves the lack of the ability to disentangle the effects of the actual interaction versus perceived supportiveness of the partner on reducing stress and worry. To examine the effect of the actual interaction, we could assess to what extent participants felt that their partners' support validated their emotion or helped them to cognitively reappraise their stressful situations. These may serve as alternative mechanisms parallel to perceived supportiveness of the partner, or act as antecedents or immediate consequences of perceived supportiveness in the causal link to reduce stress and worry. Future study is encouraged to test multiple mechanisms simultaneously.

In addition, although the single-item measure for worry has been adopted and considered valid in previous research (Rains et al., 2016), single items are more vulnerable to random measurement errors. Future research should replicate our study by using and validating a multiple-item scale for worry. Lastly, it is worthwhile for future research to either statistically control or explore the moderating effects of other factors that have been found to influence human–chatbot interactions, such as an individual's prior experience with chatbots and an individual's knowledge or belief about how AI functions.

Supporting Information

The following supporting information is available for this article:

Appendix A.

Note

1. When the hypothesis test concerns a theoretically derived hypothesis, a one-tailed hypothesis test statistically translates the logical relationship between the constructs in the hypothesis better than a two-tailed test. A two-tailed test in this case has the disadvantage of being overly conservative and inexact (Cho & Abe, 2013).

Data Availability Statement

The data underlying this article will be shared on reasonable request to the corresponding author.

Conflict of Interest

The authors do not have conflict of interest to disclose.

References

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, 42, 2355–2368.

Borkovec, T. D., Alcaine, O., & Behar, E. S. (2004). Avoidance theory of worry and generalized anxiety disorder. In R. Heimberg, D. Mennin, & C. Turk (Eds.), *Generalized anxiety disorder: Advances in research and practice* (pp. 77–108). New York, NY: Guilford.

Burleson, B. R. (2003). Emotional support skills. In J. O. Greene ?0026; & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 551–594). Mahwah, NJ: Lawrence Erlbaum.

Capobianco, L., Morris, J. A., & Wells, A. (2018). Worry and rumination: Do they prolong physiological and affective recovery from stress? *Anxiety, Stress, and Coping*, 31, 291–303.

Cho, H. C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate?. *Journal of Business Research*, 66, 1261–1266.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 385–396. https://doi.org/10.2307/2136404

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, 116, 457. 10.1037/0033-2909.116.3.457

Edwards, A., Edwards, C., Spence, P. R., Harris, C., & Gambino, A. (2016). Robots in the classroom: Differences in students' perceptions of credibility and learning between "teacher as robot" and "robot as teacher". *Computers in Human Behavior*, 65, 627–634.

Frankish, K., & Ramsey, W. M. (Eds.). (2014). *The Cambridge handbook of artificial intelligence*.Cambridge, England: Cambridge University Press.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25, 161–178. https://doi.org/10.2307/2092623

Gunthert, K. C., Cohen, L. H., & Armeli, S. (1999). The role of neuroticism in daily stress and coping. *Journal of Personality and Social Psychology*, 77, 1087–1100. https://doi.org/10.1037/0022-3514.77.5.1087

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York City, NY: Guilford Publications.

Hill, J., Randolph Ford, W., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49, 245–250. https://doi.org/10.1016/j.chb.2015.02.026

Hill, C. T., & Stull, D. E. (1982). Disclosure reciprocity: Conceptual and measurement issues. *Social Psychology Quarterly*, 45, 238–244. https://doi.org/10.2307/3033919

Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68, 712–733. https://doi.org/10.1093/joc/jqy026

Liu, B., & Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychology, Behavior, and Social Networking*, 21, 625–636. https://doi.org/10.1089/cyber.2018.0110

Jones, S. M., & Burleson, B. R. (1997). The impact of situational variables on helpers' perceptions of comforting strategies. *Communication Research*, 24, 530–555. https://doi.org/10.1177/0093650970240005004

McCormic, R. W., Pomerantz, A. M., Ro, E., & Segrist, D. J. (2019). The "me too" decision: An analog study of therapist self-disclosure of psychological problems. *Journal of Clinical Psychology*, 75, 794–800

Miner, A. S., Milstein, A., & Hancock, J. T. (2017). Talking to machines about personal mental health problems. *JAMA: The Journal of the American Medical Association*, 318, 1217–1218. https://doi.org/10.1001/jama.2017.14151

Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72, 432–440. https://doi.org/10.1016/j.chb.2017.02.067

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56, 81–103. https://doi.org/10.1111/0022-4537.00153

Priem, J. S., & Solomon, D. H. (2015). Emotional support and physiological stress recovery: The role of support matching, adequacy, and invisibility. *Communication Monographs*, 82, 88–112. https://doi.org/10.1080/03637751.2014.971416

Rains, S. A., Brunner, S. R., Akers, C., Pavlich, C. A., & Tsetsi, E. (2016). The implications of computer-mediated communication (CMC) for social support message processing and outcomes: When and why are the effects of support messages strengthened during CMC? *Human Communication Research*, 42, 553–576.

Rains, S. A., Pavlich, C. A., Tsetsi, E., & Akers, C. (2019). Evaluating the scope and theoretical mechanisms of person-centered social support: The case of the comforting computer program. *Communication Research, 47, 512–540.* https://doi.org/10.1177/0093650219852554

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Reeves, B., & Nass, C. (1996). *How people treat computers, television, and new media like real people and places.* Cambridge, England: Cambridge University Press.

Reis, H. T., Lemay, E. P., & Finkenauer, C. (2017). Toward understanding: The importance of feeling understood in relationships. *Social and Personality Psychology Compass*, 11, e12308. https://doi.org/10.1111/spc3.12308

Sarason, I. G., Sarason, B. R., Shearin, E. N. (1986). Social support as an individual difference variable: Its stability, origins, and relational aspects. *Journal of Personality and Social Psychology*, 50, 845–855. 10.1037/0022-3514.50.4.845

Shechtman, N., & Horowitz, L. M. (2003, April). Media inequality in conversation: How people behave differently when interacting with computers and people. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 5,*281–288, Ft. Lauderdale, Florida, USA.

Shenk, C. E., & Fruzzetti, A. E. (2011). The impact of validating and invalidating responses on emotional reactivity. *Journal of Social and Clinical Psychology*, 30, 163–183. https://doi.org/10.1521/jscp.2011.30.2.163

Stein, J. P., & Ohler, P. (2017). Venturing into the uncanny valley of mind: The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, 160, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger ?0026; & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (p. 73e100). Cambridge, MA: The MIT Press.

Sundar, S. S., & Kim, J. (2019). Machine heuristic: When we trust computers more than humans with our personal information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Paper No. 538). Glasgow, Scotland: ACM, May 4–9.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24–54. https://doi.org/10.1177/0261927X09351676

van der Zwaan, J. M., Dignum, V., Jonker, C. M., & van der Hof, S. (2014). On technology against cyberbullying. In *van der Hof S., van den Berg B., Schermer B. (eds)* Minding minors wandering the web: Regulating online child safety. Information Technology and Series*, 24, T. M. C. Asser Press, The Hague.*

Weiss, R. L. (1980). Strategic behavioral marital therapy: Toward a model for assessment and intervention. In J. P. Vincent (Eds.), *Advances in family intervention, assessment and theory* (Vol. 1, pp. 229–271). Greenwich, CT: JAI Press.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52, 30–41. https://doi.org/10.1207/s15327752jpa5201_2