# Ice-Breaking Technology: Robots and Computers Can Foster Meaningful Connections between Strangers through In-Person Conversations


Alex Wuqi Zhang
University of Chicago
Chicago, Illinois, USA
alexwuqizhang@uchicago.edu

Ting-Han Lin
University of Chicago
Chicago, Illinois, USA
tinghan@uchicago.edu

Xuan Zhao
Stanford University
Stanford, California, USA
xuanzhao@stanford.edu

Sarah Sebo
University of Chicago
Chicago, Illinois, USA
sarahsebo@uchicago.edu



## ABSTRACT

Despite the clear benefits that social connection offers to well-being, strangers in close physical proximity regularly ignore each other due to their tendency to underestimate the positive consequences of social connection. In a between-subjects study (N = 49 pairs, 98 participants), we investigated the effectiveness of a humanoid robot, a computer screen, and a poster at stimulating meaningful, face-to-face conversations between two strangers by posing progressively deeper questions. We found that the humanoid robot facilitator was able to elicit the greatest compliance with the deep conversation questions. Additionally, participants in conversations facilitated by either the humanoid robot or the computer screen reported greater happiness and connection to their conversation partner than those in conversations facilitated by a poster. These results suggest that technology-enabled conversation facilitators can be useful in breaking the ice between strangers, ultimately helping them develop closer connections through face-to-face conversations and thereby enhance their overall well-being.


**Figure 1: In a between-subjects study, we investigated the effectiveness of a humanoid robot, a computer screen, and a poster at stimulating meaningful, face-to-face conversations between two strangers by posing progressively deeper questions. The image shows participants in the computer screen condition discussing the question, "What is one thing you're most grateful for in your life?" The room was always equipped with a humanoid robot and a computer screen to control the presence of novel objects for all participants.**


## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**.

## KEYWORDS

Technology-Mediated Conversations, Human-Computer Interaction, Human-Robot Interaction, Social Connection

**ACM Reference Format:**

Alex Wuqi Zhang, Ting-Han Lin, Xuan Zhao, and Sarah Sebo. 2023. Ice-Breaking Technology: Robots and Computers Can Foster Meaningful Connections between Strangers through In-Person Conversations . In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3544548.3581135


## 1 INTRODUCTION

Engaging in social interactions with other people, even strangers, can have far-reaching physical and mental health benefits [32, 40, 45]. More specifically, engaging in face-to-face social interactions has been shown to be linked to greater increases in well-being than computer-mediated communication (e.g., instant messaging, social media) [28]. As the adoption of technology becomes more ubiquitous, investigating how technologies can best be leveraged to maximize well-being becomes increasingly important.

The relationship between technology and social connection has been complicated. On one hand, indiscriminately consuming sensationalized content on today's social media platforms can lead to isolation, loneliness, social comparison, anxiety, and other harmful effects [21, 56, 60]. On the other hand, technology can be instrumental in helping people forge and maintain lasting relationships [29, 31, 45], especially by enabling instant communication across vast physical barriers (e.g., Zoom, Skype, iMessage) [46]. In this study, we set out to explore whether technology can also be beneficial when used to encourage meaningful face-to-face connections between people in the same physical space.

Even without physical distance, forging meaningful connections with people can still be difficult. Recent studies have shown that despite being able to interact face-to-face, people routinely ignore strangers in their close proximity, which partly results from their



*CHI '23, April 23–28, 2023, Hamburg, Germany*



ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581135

tendency to underestimate the positive consequences of social engagement and overestimate the negative consequences [11, 12, 42], mistakenly believing that others would be uninterested in talking to them [47]. Moreover, even when people do talk to strangers, they tend to overestimate how awkward discussing personal topics might be and underestimate how interested others may be in their personal revelations, which can keep them from engaging in deeper and more meaningful conversations [26]. Given that talking about deep topics (e.g., "What is one thing you are most grateful for?") often leaves people feeling happier and more connected to their conversation partner than talking about shallower topics (e.g., "What is your favorite TV show?"), systematic miscalibrations about conversations can lead people to be less social than would be optimal for their well-being in everyday life [11, 26, 63, 64].

To explore how technology might be able to help strangers overcome the reluctance to engage in deep conversations with each other, we designed a study that leveraged different conversation facilitators to encourage pairs of unacquainted individuals to converse with each other. We investigated three different conversation facilitators that vary in their intrusiveness and human-likeness: a humanoid robot, a computer screen, and a poster. We found that participants in conversations facilitated by either the humanoid robot or the computer screen reported greater happiness after the experience as well as a higher level of connectedness to their conversation partner than those in conversations facilitated by a poster displaying the same question prompts. As technology becomes an essential part of our social lives, our work opens new possibilities for technologies to facilitate in-person social interactions and deep conversations for greater connectedness and emotional well-being.

## 2 BACKGROUND

In this section, we review literature related to how strangers engage in conversations, including the barriers they experience when starting conversations. We also examine prior work investigating the current state of computer-mediated conversations in groups of people who have not previously met before. Noting the unique social attributes of humanoid robots and intrigued by the role they could potentially play in helping strangers break the ice, we also reviewed work exploring how robots can shape conversations between people.

### 2.1 Conversations between Strangers

An average person's daily routine often includes many opportunities for social interaction with strangers, an activity that increases not only happiness and well-being [17, 32, 40, 44, 45] but also knowledge bases [4]. Compared to other forms of socialization, such as talking over the phone or sending an email, these in-person social occasions give people the advantage of having face-to-face interactions with another person at the same time and in the space. However, in the absence of a strong norm for face-to-face socialization, people often readily forgo these opportunities to talk to strangers and form meaningful connections with them even if doing so can improve their emotional well-being [11] and generate informational benefits [4]. For instance, multiple studies have shown that people tend to underestimate how positively they would feel after talking with strangers in close physical proximity, such as on a commuter train [12, 47]. Obstacles preventing people from conversing with strangers partly stem from the fact that they are uncertain whether the other person will be interested in talking to them, whether it is socially appropriate to start a conversation, or even what they should talk about [11, 42]. Even if individuals overcame these psychological barriers and actually started a conversation with a stranger, two unacquainted interlocutors are more likely to converse over a shallow, neutral topic like the weather because they want to appear more polite and less intrusive to the other person [53].

To overcome the barriers people experience when starting conversations with strangers, psychologists have suggested several solutions to help people forge meaningful connections with one another. Some researchers designed a series of scavenger hunts that involved "repeatedly finding, approaching, and talking to strangers" and found that this intervention program improved people's outlook on connecting with strangers [43], whereas others have focused on helping people identify meaningful conversation topics to reap the benefits of self-disclosure. For instance, Aron et al.'s famous 36 personal questions, publicized in The New York Times [22], is a fast-friend paradigm that elicits self-disclosure between strangers as "a practical methodology for creating closeness" [3]. More recently, Kardas et al. provided a list of deep and shallow discussion questions to pairs of strangers, and across a dozen experiments, they found that pairs who discussed the deep questions (e.g., "For what in your life do you feel most grateful?") felt happier, more connected, and less awkward than they initially expected [26]. In our study, we explore how different forms of technology can be best leveraged to lower the psychological barriers of social connection and encourage meaningful face-to-face conversations.

### 2.2 Computer-Mediated Conversations between Strangers

Technology has been frequently used to facilitate conversations between two strangers in many computer-mediated scenarios. Prominent examples of such technology includes online dating and online chatting. Based on a survey conducted in 2019, Anderson et al. reported that 23% of Americans have had a long-term relationship with another person they met on a dating app [1]. Through the matching process in dating apps, users can confirm their partner's interest and thus start a conversation [19]. Another popular computer-mediated venue used to talk to strangers is anonymous online chatting. These online chat rooms offer anonymity for users, which in turn lowers the stakes associated with engaging in unfiltered or intimate conversation with strangers [7].

Additionally, Shin et al. found that a machine learning powered chatbot that suggests personalized topics and questions while people chat with each other in an online chat room can increase conversation quality as well as perceived closeness between conversation partners [50]. In face-to-face conversations, Yang et al. built a distributed mobile communications system that suggests conversation topics via Bluetooth Service Discovery Protocol [61] and Nguyen et al. found that using Google Glass to deliver real-time personalized discussion topics (e.g., a mutual interest or hobby) to a stranger pair through a ranking recommendation algorithm could strengthen participants' interpersonal interaction [35].

While these technology-mediated methods of encouraging conversations between strangers have demonstrated that technology can positively influence the conversation experience between unacquainted individuals, there have not been many technologies developed for actively prompting people to engage in face-to-face, in-person conversations with strangers. However, we believe that solutions may be "hidden" in plain sight in our everyday environments. For example, public spaces often use digital screens to display commercial content or information for public benefits (e.g., personal hygiene practices in a doctor's office). Could we use those monitors to display conversation prompts that have been shown to foster more meaningful connections among strangers? Given their prevalence in public spaces, computer screens in public spaces may present missed opportunities to facilitate better social connections among strangers.

## 2.3 Robot-Mediated Conversations and Interactions between People

Prior work in human-robot interaction (HRI) suggests that robots can shape conversations and interactions between people (see [48] for a review). Robots can exert influence on a group's conversation and moderate participation among its members by producing organic gestures [13, 27], orienting their gaze [15, 34], and making supportive comments [51]. When operating as part of a human-robot team, a robot that expresses vulnerable self-disclosures has been shown to both increase the number of vulnerable expressions by the human team members [52] and encourage greater conversations between the human team members [54]. Robots have also demonstrated their effectiveness in successfully mediating conflict between children [49] and increasing a team's awareness of an ongoing conflict [23]. Finally, robots have shown promise in emerging use cases that can be highly personal, such as facilitating couples' counseling sessions that have resulted in increased intimacy and positivity between participating couples [58].

We chose a humanoid robot as one of the conversation facilitators because of the range of human behaviors it can influence and demographics it has been shown to be able to impact, as well as its demonstrated ability to elicit greater compliance from human users compared to purely video-based agents [5]. No work to our knowledge has investigated a robot's ability to engage two strangers in a conversation for the purpose of forming interpersonal connections, nor how robots compare to other technologies in this undertaking.

# 3 METHODS

In a between-subjects study, we recruited 49 pairs of participants (98 individuals total) who were unacquainted with each other to sit for 10 minutes in a waiting room under the cover story of a short-term memory retention challenge. While in the room, a device serving as a conversation facilitator—a humanoid robot, a computer screen slideshow, or a poster on the wall—would present a pair with progressively deeper and more intimate questions to discuss with each other. We video-recorded these interactions for behavioral measures and collected participants' self-reported survey responses before and after their conversations. The study protocol was approved by the University of Chicago's Institutional Review Board (IRB20-1951).

## 3.1 Hypotheses

In this study, we investigated three different conversation facilitators (humanoid robot, computer screen, and poster) and their ability to alleviate the psychological barriers preventing strangers from engaging in conversation with each other. Of the different conversation facilitators in our study, we predicted that the facilitators with greater social presence would be more effective at breaking the ice and getting the strangers to interact with each other. Since prior work has demonstrated that physically embodied social robots command greater social presence than purely video-based agents [5, 24], we hypothesized:

- **$H_1$ – More participants will speak with the other participant** while they are in the waiting room when the conversation facilitator is a robot, then a computer screen, then a poster, in that order.
- **$H_2$** – Participants will think it was **easier to start a conversation** and feel less awkward and vulnerable while starting the conversation when the conversation is facilitated by a robot, then a computer screen, then a poster, in that order.

As the conversation questions being posed got progressively more personal, we anticipated that participants would experience a higher level of vulnerability and awkwardness [6, 30] during their conversation experience, especially when asked to discuss deeper questions. Naturally, the escalated vulnerability and awkwardness may cause them to be less inclined to discuss the deep questions. It is possible, however, that technological conversation facilitators could induce greater compliance in participants responding to the conversation questions they pose, thereby helping participants overcome their natural inclination to avoid discussing the deep questions. Robots, in particular, may be likely to create the greatest participant compliance, since prior work has demonstrated that people are more willing to follow the instructions of a physically embodied robot as opposed to a video of that same robot displayed on a computer screen [5]. As such, we predicted that:

- **$H_3$** – Participants will **spend more time discussing the deep questions** when the conversation is facilitated by the robot, followed by the computer screen and then the poster, in that order.

Prior work in psychology has also identified positive correlations between vulnerable self-disclosure and interpersonal closeness [3, 10]. Furthermore, studies have also suggested that deeper conversations, as opposed to small talk, are more effective at eliciting positive emotions and interpersonal closeness [3, 33]. Because the progressively deeper conversation questions in our study were specifically designed to stimulate more personal discussions and encourage more self-disclosure, we hypothesized that:

- **$H_4$** – Participants in the robot condition will experience the highest level of **interpersonal closeness**, followed by those in the computer screen and poster conditions, in that order.
- **$H_5$** – Participants in the robot condition will have the most **positive conversation experience**, followed by the computer screen and poster conditions, in that order.

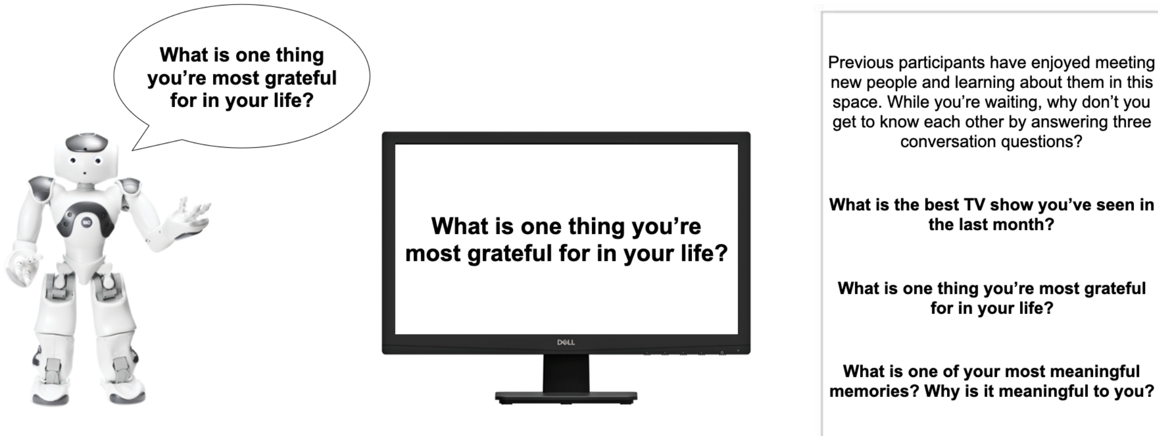


**Figure 2: The study investigated the effectiveness of three different conversation facilitators: a humanoid robot (left), a computer screen (middle), and a poster (right). The questions posed in each condition were identical.**

## 3.2 Experimental Conditions

The experiment included three between-subjects conditions that corresponded to the type of device facilitating the conversation: a robot condition, a computer screen condition, and a poster condition (see Figure 2). The poster was chosen to be a baseline conversation facilitator as it is represents a common and low-cost method of information delivery in public spaces. Like posters, digital displays are also frequently utilized in public places to present information, which informed our choice of a computer screen conversation facilitator that displays dynamic content. Finally, we elected to have a humanoid robot serve as the third conversation facilitator because they have demonstrated unique advantages in eliciting greater compliance from people in social interactions [5], which we hypothesize may encourage strangers to talk with one another more effectively than the other conversation facilitators. As the content the conversation facilitators were delivering was social in nature, we believed it would be beneficial to use a human-like robot that was able to convey this content using similar social signals to the ones humans use (e.g., eye gaze, gestures).

In the **robot condition**, an experimenter controlled a Softbank Robotics NAO robot using the Wizard of Oz paradigm [41]. The humanoid robot introduced itself and posed three questions to the participants, pausing for approximately three minutes after each question for the participants to discuss it together. We selected a NAO robot to facilitate the conversation because it was designed to interact with people through speech and gestures. In order to standardize the range of manipulations across all three conditions, we programmed the robot to not respond to participants unless they asked it to repeat a question. After the interaction concluded, the robot thanked the participants for sharing their thoughts with each other.

In the **computer screen condition**, a monitor with a screen 24 inches wide by 14 inches high displayed one of the three questions for approximately three minutes. The monitor's (slideshow) transitions from question to question were complemented with a sound effect of a ringing bell, alerting participants to the new question. We selected a computer screen to be one of the conversation facilitators due to its programmable visual interface as well as the fact that it is highly scalable and easy to implement in the real world.

In the **poster condition**, a 24 inch by 36 inch poster prompted participants to engage in a conversation using a list of three questions. The poster was fixed to the wall behind the computer screen and the humanoid robot in a way that it was not occluded from the participants' line of sight. The poster was chosen to be a baseline condition where the same information was delivered to participants without assistance from technological devices. Finally, unlike the robot and monitor, which were always present in the room regardless of conditions (but were turned off when not in use), the poster was only present in the poster condition.

## 3.3 Conversation Questions

In all three conditions, the conversation facilitator prompted the participants to engage in a conversation with the same set of questions (see Figure 2), presented in the following order:

(1) What is the best TV show you've seen in the last month?
(2) What is one thing you're most grateful for in your life?
(3) What is one of your most meaningful memories? Why is it meaningful to you?

The order of the questions was meant to ease the transition from a shallow question (1) to the deeper questions (2, 3). We selected and modified these questions from the 36 personal questions used in Aron et al. [3] to examine human connection between strangers.

## 3.4 Cover Story

In our study, we used a cover story in the form of a short-term memory challenge to better measure how participants would respond to prompts for conversation without any prior expectations that they would be talking with a stranger. At the time they were recruited, participants were told that they were participating in a memory challenge that required a waiting time before they were tested on their information retention.

To avoid priming participants with the true purpose of our study, we informed participants that the study centered on short-term memory retention. Participants began by reading an article about a neutral topic (externalities, self-serving bias, measuring well-being, or working memory). Then, they were told to spend 10 minutes in a waiting room before they would be tested on how much of the article they could recall. During their 10 minutes in the waiting room, we conducted our actual experiment, where a conversation facilitator (i.e., humanoid robot, computer screen, or poster) attempted to get the two participants in the waiting room to engage in conversation. After spending 10 minutes in the waiting room, participants answered a series of questions about the article in part of the post-experiment survey. Upon their completion of the post-experiment survey, the participants were debriefed by an experimenter about the true purpose of the study and were also given the opportunity to withdraw their data from the study. No participant to our knowledge anticipated the true purpose of the study during the experiment. While a few participants speculated whether the conversation facilitators were trying to "distract them from the memory challenge" and "get them to forget the passage they read earlier", we do not believe participants saw through the cover story and altered their behavior because of it.

## 3.5 Study Protocol

The study was conducted at Mindworks, a behavioral science center open to the public in downtown Chicago, over a period of approximately three months, from February 19$^{th}$, 2022 to May 1$^{st}$, 2022. We mostly conducted experiments on Fridays, Saturdays, and Sundays, at a rate of 3 − 5 pairs per day. Individuals walking by the center were recruited to participate in a "Memory Challenge" study. The research assistants conducting the experiment recruited participants from distinct groups who did not know each other before they started the experiment. Once two participant candidates became available around the same time, they were prompted to follow a researcher into a private room and were seated in front of two computers that were placed back-to-back and separated by a divider. After consenting, participants completed a Qualtrics survey featuring a short article on a neutral topic (i.e., the cover story), followed by a question on their current mood on a scale from −5 (*very negative*) to +5 (*very positive*). Once both participants finished the pre-conversation questionnaire, the researcher guided them to an adjacent waiting room approximately 7ft by 7ft in size.

Next, the researcher instructed the participants to remain seated in the room for 10 minutes before returning to the computer room to test how much information they retained about the article. Regardless of the experimental condition, the waiting room was always equipped with a humanoid robot and a monitor to hold the presence of novel objects in the room constant for all participants (see Figure 1). Because the robot itself could be an interesting topic of conversation, we controlled for the possibility that the presence of these devices might influence the participants' conversation experience (e.g., all participants would be equally likely to discuss the appearance of the robot). A poster was only present in the waiting room during the poster condition. Once the two participants were alone in the waiting room, conversation questions were posed to the participants by either a robot, a computer screen, or a poster, depending on the condition each pair was randomly assigned into.

In the robot condition and the computer screen condition, another researcher (not visible to the participants) would control the robot and the monitor to pose the conversation questions. The technological facilitators posed the first conversation question as soon as the two participants were left alone. If the conversation continued past 3 minutes or the conversation stalled for 10 seconds, the technological facilitator would interrupt the participants to pose the next conversation question.

After the participants spent 10 minutes in the waiting room, a researcher would re-enter the waiting room to announce that the waiting period was over and guide the participants back to the initial computer room, where they answered questions related to the short article they had read as well as their conversation experience in the waiting room.

After completing the post-conversation survey, each participant was debriefed on the true objective of the study. The study took approximately 20 minutes, and all participants were given compensation equivalent to $4.00.

## 3.6 Measures

To assess how effective the devices were at facilitating deep conversations between human participants, we examined participants' behavior in the waiting room as well as their responses to the post-conversation questionnaire. Throughout the questionnaire, we used scales of different ranges based on the original sources of the scales we referenced in order to make our work more easily comparable to other work.

*3.6.1 Demographics and Personality Attributes.* We gathered general demographic data about participants including age, ethnicity, and gender. Additionally, we assessed participants' personality on the dimensions of extraversion and openness to new experiences through the Ten Item Personality Measure (TIPI) [16] on a 7-point Likert scale from 1 (*disagree strongly*) to 7 (*agree strongly*).

*3.6.2 Starting the Conversation.* If the participants talked to each other during their time in the waiting room, we asked them to rate how awkward, uncomfortable, easy, and vulnerable they felt when starting the conversation (e.g., "How **awkward** was it to start the conversation?") on scales from 1 (*not at all*) to 10 (*very much so*).

*3.6.3 Conversation Experience.* We used scales ranging from 1 (*not at all*) to 10 (*very much so*) to examine the participants' conversation experience. The majority of the questions were adapted from Kardas et al. [26], which included: "How **happy** do you feel about your conversation?", "How **awkward** was the conversation?", "How **uncomfortable** was the conversation?", "How **pleasant** was the conversation?", "How **enjoyable** was the conversation?", and "How **vulnerable** did you feel during the conversation?" In addition, participants also reported how their interest in engaging in a conversation with people they did not know had changed after the experience in the waiting room using a scale that ranged from −5 (*decreased greatly*) to +5 (*increased greatly*). Finally, participants reported their mood after their experience in the waiting room on the same scale ranging from −5 (*very negative*) to +5 (*very positive*).

*3.6.4 Social Perception of the Other Participant.* To directly probe what participants thought of their conversation partner, we asked them to respond to four scales each ranging from 1 (*not at all*) to 10 (*very much so*) : "How **connected** do you feel to your conversation partner?", "How much do you **like** your conversation partner?", "How well do you feel you got to **know** your conversation partner?" and "How much did you **learn** from your conversation partner?"

*3.6.5 Facilitator Social Attributes.* As a means to assess the conversation facilitator's warmth and sincerity, we used the "warmth" sub-scale from the Robotic Social Attribute Scale (RoSAS) [8] and the "sincerity" sub-scale from the Multi-Dimensional Measure of Trust (MDMT) [57] respectively. Both were 10-point scales from 1 (*definitely not associated*) to 10 (*definitely associated*). Additionally, we included 7-point questions from 1 (*strongly disagree*) to 7 (*strongly agree*) derived from the "Co-presence" sub-scale from the Social Presence Scale [18] to measure the degree of social facilitation the conversation facilitator had on the participants. While there is no clear consensus on the definition of social presence, we find it most appropriate for this study to define social presence as the degree to which a person's perceptions of an agent shape social interaction with that agent [5].

*3.6.6 Facilitator Usefulness.* To measure how useful people believed the conversation facilitators could be, we introduced three different scenarios, including Networking Session, Grocery Store, and Home Use to probe for participant intuitions (e.g., "The robot could be useful in a **networking session** where you do not know the majority of the attendees."). For each scenario, participants responded on a scale from 1 (*strongly disagree*) to 10 (*strongly agree*).

*3.6.7 Interactive IOS Scale for Multiparty Interactions: A Measure of Interpersonal Closeness with Multiple Interactants.* We prompted participants to report their perceived interpersonal closeness by arranging three circles (representing the participant, the conversation partner, and the facilitator) via an interactive web interface we developed for this study. As shown in Figure 3, the three circles were initially positioned in a diagonal, and participants were asked to "drag the circles into the arrangement that best represents your relationship with the other participant and the [Robot / Computer Screen / Poster]." This measure is adapted from Aron's Inclusion of Other in the Self Scale [2] to capture people's perception of social closeness among multiple interactants, and it is also grounded in a wealth of psychological studies that measured social distance based on physical distance between individuals [20] as well as their computerized protagonists [39]. Hence, we also measured the interpersonal closeness of the two participants using the distance (in pixels) between the center of the "You" circle and that of the "Other Participant" circle. The JavaScript code and instructions for how to use this scale can be found here[1].

*3.6.8 Behavioral Measures: Conversation Question Compliance.* Three independent coders watched videos of the participants in the waiting room and marked the time participants spent talking and whether that time was spent discussing the conversation questions posed by the facilitator. The coders were in agreement 95% of the time.

[1]https://github.com/SeboLab/interactive_ios_scale

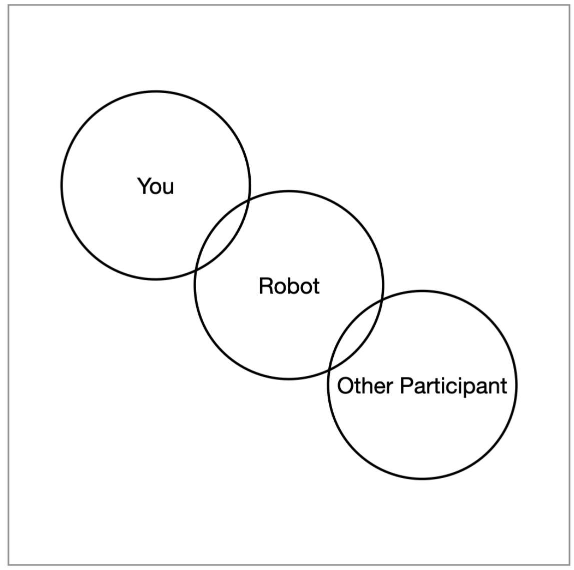


**Figure 3: We developed the Interactive IOS Scale for Multiparty Interactions measure to assess the level of connection participants perceived between themselves, the other participant, and the conversation facilitator (robot, computer screen, or poster). We have open sourced this measure to enable other researchers to adapt it to their use cases.**

This agreement percentage was calculated by selecting an overlap set of 5 videos (11% of the total data) and computing the percentage in which labeled time intervals intersected across the three coders.

Based on the coding results, we computed the following behavioral measures:

- *Total Talking Percentage*: Percentage of the total duration spent in the waiting room that participants spent talking.
- *On Topic Percentage*: Percentage of the total talking time that participants spent discussing the questions posed by the conversation facilitator.
- *Deep Question Compliance Percentage*: Percentage of the total talking time that participants spent discussing the **two deep questions** posed by the conversation facilitator. This measure was denoted by $P$, with $Q2$ and $Q3$ representing the amount of time spent discussing question 2 and question 3, respectively, and $Total\_Talk$ representing the total amount of talking time.

$$P = \left( \frac{Q2 + Q3}{Total\_Talk} \right)$$

*3.6.9 Content Analysis on Conversation Depth.* To supplement the Conversation Question Compliance metric, we analyzed the content of the conversations with both human coding and dictionary-based computerized coding program. We had three independent coders review the recordings of the participants' conversations and assigned a numerical *depth rating* between 1 to 5. The coders followed the following scheme when assigning their ratings:

- Depth rating of 1: Less than 20% of the conversation is spent discussing personal topics (including but not limited to the questions posed by the conversation facilitator)
- Depth rating of 3: About 50% of the conversation is spent discussing personal topics
- Depth rating of 5: More than 80% of the conversation is spent discussing personal topics

Examples of personal topics included marriage, immigration, and major career decisions. The coders were in agreement 94% of the time. This agreement percentage was calculated by selecting an overlap set of 6 videos (12% of the total data) and computing the average of the absolute difference across the 6 videos.

In addition, we transcribed all conversations using Otter.ai and analyzed them with Linguistic Inquiry and Word Count ("LIWC") [37], a rigorously developed and widely adopted text analysis program that can calculate the percentage of words in given linguistic, psychological and topical categories. Given our research questions, we were particularly interested in examining whether conversations across different conditions involved different amount of personal disclosure and sense of connection.

*3.6.10 Coding of Free Response Questions.* To get a more holistic understanding of the participants' experience during the conversation and with the conversation facilitator, we included several free response questions in the post-experiment survey: "Please briefly describe what happened during the time you just spent in the waiting room" and "How would you best describe the role of the [robot / computer monitor / poster] in the conversation?" We then had two independent coders (Cohen's kappa coefficient $\kappa > 0.85$) review the participants' responses and label their entries with whether or not they belong to certain predetermined categories. The categories we used for the question pertaining to the time participants spent in the waiting room included "answered questions," "discussed personal topics," "discussed the study," and "discussed miscellaneous shallow topics." The categories we used for the question about the participants' perception of the conversation facilitator's role included "facilitator," "involved," and "no role."

## 3.7 Participants

A total of 49 pairs (98 individuals) participated in the study. Of the 49 pairs, we excluded 5 pairs from our analysis due to technical difficulties (e.g., robot or computer screen lost connection). Of the remaining 44 participant pairs (88 individuals) included in our data analysis, 14 were in the robot condition, 15 were in the computer screen condition, and 15 were in the poster condition.

Of the 88 participants, 3 did not disclose their demographic data. The remaining 85 participants ranged in age from 18 to 65 ($M = 32.47$, $SD = 10.96$). 50 participants identified as female, 33 as male, and 2 as non-binary.

Among the participants, 48 identified as White, 12 as East Asian, 8 as South Asian, 12 as Hispanic, 5 as Black, 2 as Middle Eastern, and 3 as Other. Those who identified themselves as multiple ethnicities were double counted in those ethnicities.

Participants reported their extraversion ($M = 4.36$, $SD = 0.78$) and openness to new experiences ($M = 4.53$, $SD = 0.82$) on a 7-point scale. No statistically significant differences were found across the three conditions on personality traits.

# 4 RESULTS

To analyze our data where each participant was represented as a single data point, we used linear mixed-effects models (LMM) to account for participants being grouped into conversation pairs. In our LMMs, we assigned the condition (robot, computer screen, or poster) as a fixed effect and designated conversation pair as a random effect. We examined the residuals in our LMMs and evaluated residual errors for lack of trends and heteroscedasticity. We reported the fixed effect's linear coefficient ($\beta$), standard error ($SE$), and $p$ value from each LMM. The sign (positive, negative) of the linear coefficient ($\beta$) is dependent on the direction of comparison between conditions, i.e., which condition is the reference condition. We used the poster condition as the reference condition in our LMMs when comparing the poster condition to the robot and computer screen conditions. When comparing the robot condition to the computer screen condition, we used the robot condition as the reference condition.

For pair-level data where each conversation pair is represented as a single data point, we used analysis of variance (ANOVA) tests to examine the effect of the experimental conditions on the pair. We then conducted Tukey's Honest Significant Differences tests and reported the $p$ value for the pairwise-comparison among the three conditions.

## 4.1 Social Presence and the Role of the Conversation Facilitator

As a manipulation check to ensure that participants viewed the conversation facilitators differently as we intended, we first analyzed the perceived social presence of each conversation facilitator. We anticipated that the robot would have the greatest perceived social presence, then the computer screen, and finally, the poster. We did find that participants reported the robot ($M = 5.02$, $SD = 1.22$) as having a significantly higher social presence than the poster ($M = 3.55$, $SD = 1.32$, $\beta = +1.47$, $SE = 0.33$, $p < 0.001$) and a marginally higher social presence than the computer screen ($M = 4.40$, $SD = 0.98$, $\beta = -0.62$, $SE = 0.33$, $p = 0.070$). Participants also reported the computer screen as having a greater social presence than the poster ($\beta = +0.85$, $SE = 0.33$, $p = 0.013$).

To gauge how participants perceived the role played by the different conversation facilitators, we further examined their responses to the open-ended question, "How would you best describe the role of the [robot / computer monitor / poster] in the conversation?" Participants in the robot condition ($M = 32.14\%$, $SD = 47.56$) were significantly more likely to mention that they saw the robot as a "facilitator" than participants in the poster condition ($M = 10.00\%$, $SD = 30.51$, $\beta = 22.14$, $SE = 10.55$, $p = 0.042$). We did not observe significant differences between the computer screen condition and the robot ($p = 0.256$) or the poster condition ($p = 0.340$).

Conversely, significantly more participants in the poster condition ($M = 36.67\%$, $SD = 49.01$) mentioned that they perceived the poster to have "no role" in their conversation experience compared to participants in the robot condition ($M = 0.00\%$, $SD = 0.00$, $\beta = -36.67$, $SE = 11.80$, $p = 0.003$). No significant differences were observed between the computer screen condition and the other two conditions.

Overall, these results confirmed that participants' experience with the conversation facilitators was consistent with our intention of the study, and the robot stood out as having the greatest social presence. However, it is important to note that the three conversation facilitators likely differed on more dimensions than just social presence.

## 4.2 Starting a Conversation with Strangers

Our first hypothesis ($\mathbf{H}_1$) predicted that more participants in the robot condition would engage in conversation with the other participant in the waiting room than those in the computer screen condition, who would in turn engage more than those in the poster condition. Interestingly, every pair of participants ended up starting a conversation with each other across all conditions. This is likely due to the fact that participants who visited the behavioral science center and volunteered to take a novel study were already in a relatively exploratory mindset that deviated from everyday routines. They were also guided into the waiting room simultaneously and were the only two people in that small room. Despite our intention to stage a waiting room scenario to mimic real life, it is evident that our setup still differed considerably from real-world scenarios (e.g., doctor's office waiting rooms, public buses), where people hold very different expectations about having conversations in public settings. Since we did not see any differences in the number of participant pairs that talked, $\mathbf{H}_1$ is not supported.

To address $\mathbf{H}_2$, which pertained to the participants' thoughts on their experience starting the conversation with someone they had never met before, we examined participants' ratings of how awkward, uncomfortable, easy, and vulnerable they felt at the beginning of the conversation. We did not find any significant differences in participant ratings. Overall, on a scale of 1 (*not at all*) to 10 (*very much so*), participants reported it being easy to start a conversation across all three conditions: robot ($M$ = 7.46, $SD$ = 2.33), computer screen ($M$ = 7.20, $SD$ = 2.37), and poster ($M$ = 6.70, $SD$ = 2.71). They also felt somewhat vulnerable across all three conditions: robot ($M$ = 4.14, $SD$ = 2.24), computer screen ($M$ = 4.50, $SD$ = 2.71), and poster ($M$ = 3.20, $SD$ = 2.09). It is possible that the type of conversation facilitator had no influence on how participants perceived the start of the conversation. It is also possible, however, that the ensuing conversation overshadowed the brief amount of time participants spent actually getting the conversation started. Regardless, $\mathbf{H}_2$ is not supported by our data.

## 4.3 Conversation Question Compliance

To test $\mathbf{H}_3$, which pertained to each facilitator's ability to focus participants' conversation around the two deep questions, we analyzed participants' actual conversation behaviors (see Figure 4) on the following metrics.

For *Total Talking Percentage*, we did not notice any significant differences between the robot ($M$ = 92.41%, $SD$ = 10.09), computer screen ($M$ = 95.32%, $SD$ = 8.56), and poster ($M$ = 90.13%, $SD$ = 18.93) conditions.

For *On Topic Percentage*, participants in the robot condition ($M$ = 48.11%, $SD$ = 26.24) spent a significantly greater percentage of time talking about the conversation questions than those in the poster condition ($M$ = 27.00%, $SD$ = 21.77, $F$ = 3.21, $p$ = 0.040, $\eta^2$ = 0.14),

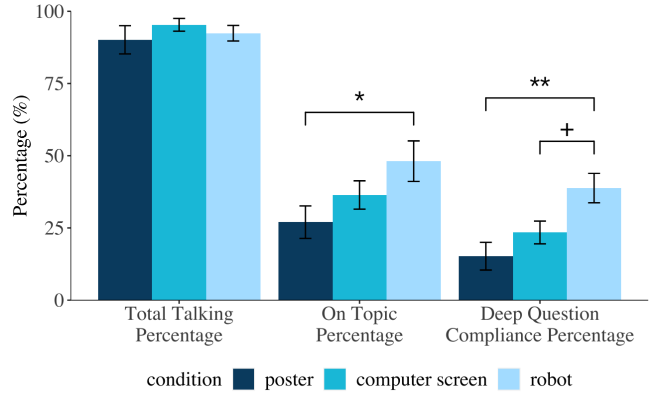


**Figure 4: Participants spent more time discussing the deep conversation questions in the robot condition compared to the poster condition. (**) denotes $p < 0.01$, (*) denotes $p < 0.05$, and (+) denotes $p < 0.10$. Error bars depict one standard error from the mean.**

as shown in Figure 4. However, the difference between the robot condition and the computer screen condition was non-significant, despite it being in the predicted direction ($M$ = 36.41%, $SD$ = 18.99, $p$ = 0.348). Similarly, participants in the computer screen condition had a higher *On Topic Percentage* than those in the poster condition ($p$ = 0.490), but the difference was not statistically significant.

Next, we examined the *Deep Question Compliance Percentage*, the metric that provides the most direct answer to $\mathbf{H}_3$, pertaining to the expected difference in participants' compliance with discussing the two deep conversation questions posed by the facilitator. As expected, participants in the robot condition ($M$ = 38.81%, $SD$ = 19.06) had the highest *Deep Question Compliance Percentage*, significantly more than those in the poster condition ($M$ = 15.23%, $SD$ = 18.57, $F$ = 6.60, $p$ = 0.002, $\eta^2$ = 0.24) and marginally more than those the computer screen condition ($M$ = 23.43%, $SD$ = 15.26, $p$ = 0.061), as shown in Figure 4. There was no significant difference between the computer screen condition and the poster condition ($p$ = 0.419). As such, $\mathbf{H}_3$ is partially supported.

In addition to our primary measures, we also examined participants' responses to the question "Please briefly describe what happened during the time you just spent in the waiting room." Consistent with results above, significantly more participants in the robot condition ($M$ = 82.14%, $SD$ = 39.00) mentioned that they "answered the conversation questions" than in the poster condition ($M$ = 46.67%, $SD$ = 50.74, $\beta$ = +35.48, $SE$ = 13.45, $p$ = 0.012). There were no significant differences between the computer screen condition and the robot ($p$ = 0.257) or the poster condition ($p$ = 0.138).

## 4.4 Participant Interpersonal Closeness

To address $\mathbf{H}_4$, which predicted that participants in the robot condition would feel the closest to their conversation partners, we first examined participants' responses to a series of statements related to how they viewed their conversation partner. We did not see any significant differences between the three conditions for any of these measures.

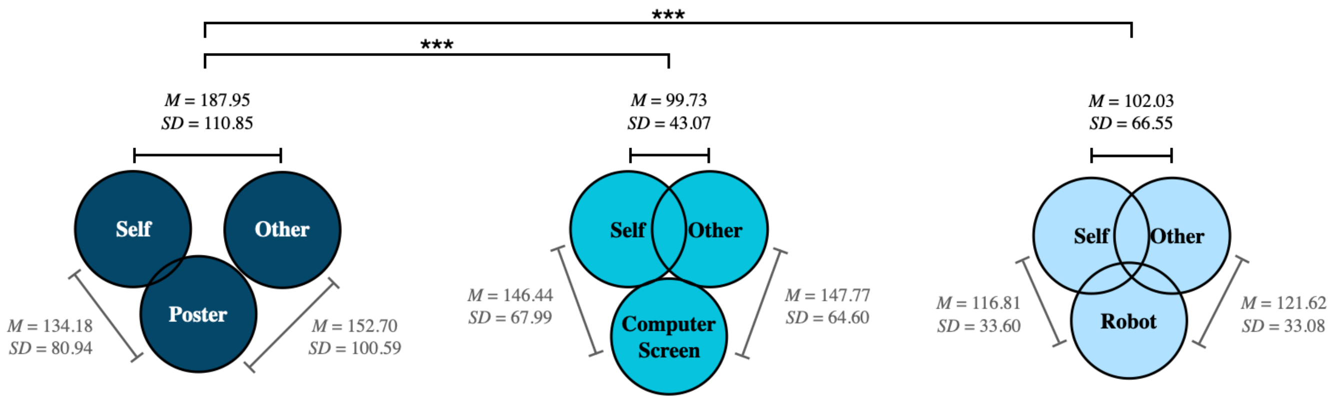


**Figure 5: The length of the lines between the centers of the circles denote the average perceived distance (in pixels) between the participant themselves, their conversation partner, and the facilitating device. When participants were asked to rearrange the circles to most accurately represent the relationship between the three entities, the distance between themselves and their conversation partner was closer when the facilitator was a robot or a computer screen, as opposed to a poster. (***) denotes $p < 0.001$ and the radius of each circle is 75 pixels.**

However, participants' responses to the *Interactive IOS Scale for Multiparty Interactions* measure revealed significant differences across conditions, likely because this measure is both more precise (arranging circles offers more granularity) and more accurate due to the presence of a neutral reference point – the conversation facilitator (see Figure 5). We found that participants indicated a significantly greater distance between themselves and their conversation partner in the poster condition ($M = 187.95$ pixels, $SD = 110.85$) than in both the robot condition ($M = 102.03$ pixels, $SD = 66.55$, $\beta = -85.92$, $SE = 20.74$, $p < 0.001$) and the computer screen condition ($M = 99.73$ pixels, $SD = 43.07$, $\beta = -88.22$, $SE = 20.38$, $p < 0.001$). By contrast, there were no significant differences between the computer screen condition and the robot condition ($\beta = -2.30$, $SE = 20.74$, $p = 0.912$). Overall, results from our *Interactive IOS Scale for Multiparty Interactions* measure demonstrated that participants felt closer to their conversation partner when the conversation was facilitated by a robot or a computer screen, as opposed to a poster. We conclude that $\mathbf{H}_4$ is partially supported.

## 4.5 Conversation Experience

To investigate $\mathbf{H}_5$, pertaining to participants' enjoyment of the conversation based on the conversation facilitator, we examined how participants responded in the post-conversation questionnaire about their conversation experience: awkward, uncomfortable, vulnerable, pleasant, enjoyable, and happy. We observed significant differences between our experimental conditions for how *vulnerable* participants felt during the conversation and how *happy* participants felt about the conversation (see Figure 6). We also saw marginally significant differences for how *enjoyable* participants felt the conversation was.

Participants in the poster condition ($M = 7.40$, $SD = 2.16$) reported that the conversation experience was marginally less enjoyable than those in the computer screen condition ($M = 8.23$, $SD = 1.25$, $\beta = +0.83$, $SE = 0.48$, $p = 0.088$). There were no significant differences between the robot ($M = 8.21$, $SD = 2.08$) and the poster condition nor between the robot and the poster condition.

Participants in the poster condition ($M = 6.96$, $SD = 2.30$) reported feeling significantly less happy about the conversation than both participants in the robot condition ($M = 8.14$, $SD = 1.82$, $\beta = +1.18$, $SE = 0.49$, $p = 0.021$) and the computer screen condition ($M = 7.97$, $SD = 1.25$, $\beta = +1.00$, $SE = 0.48$, $p = 0.044$). There was no significant difference in the happiness participants reported feeling about the conversation between the robot condition and the computer screen condition. Therefore, we conclude that $\mathbf{H}_5$ is partially supported. It is interesting to note that participants who felt the most vulnerable also reported feeling the happiest.

Furthermore, participants in the robot condition ($M = 5.21$, $SD = 2.87$) reported feeling significantly more vulnerable during the conversation than those in the poster condition ($M = 2.93$, $SD = 2.32$, $\beta = +2.28$, $SE = 0.74$, $p = 0.004$). Those in the computer screen condition did not report feeling significantly more or less vulnerable ($M = 4.00$, $SD = 2.56$) than participants in the robot ($p = 0.107$) or poster conditions ($p = 0.148$).

Additionally, when analyzing participants' change in mood before and after the experiment, we noticed significant mood improvements among participants within the robot and computer screen conditions. More specifically, participants in the robot condition reported feeling slightly negative before the conversation experience ($M = -0.75$, $SD = 3.90$) and felt significantly more positive after the conversation ($M = +1.25$, $SD = 2.08$, $\beta = -2.00$, $SE = 0.76$, $p = 0.012$). Similarly, participants in the computer screen condition started with a slightly negative mood ($M = -0.77$, $SD = 3.92$) and ended up feeling more positive after the conversation ($M = +1.30$, $SD = 2.18$, $\beta = -2.07$, $SE = 0.76$, $p = 0.010$).

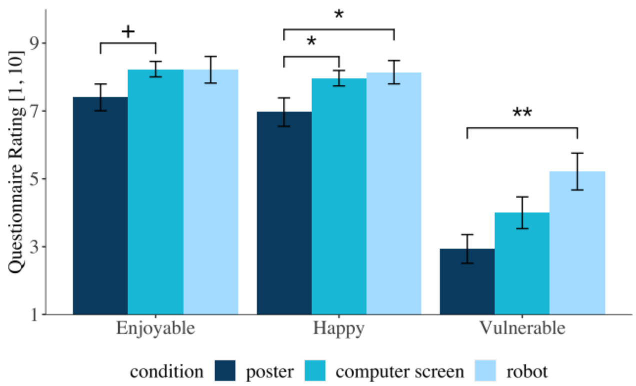


**Figure 6: Participants rated their conversation experience in terms of how enjoyable the conversation was, how happy they felt about the conversation, and how vulnerable they felt during the conversation. (**) denotes $p < 0.01$, (*) denotes $p < 0.05$, and (+) denotes $p < 0.10$. Error bars depict one standard error from the mean.**

### 4.6 Supplemental Analysis on Conversation Depth

Building onto our *Deep Question Compliance Percentage* reported in Section 4.3, we analyzed the depth ratings assigned by independent coders and found that conversations in the robot condition ($M$ = 3.38, $SD$ = 0.87) were viewed as significantly deeper (i.e., participants spent more time discussing personal topics) than those taken place in the poster condition ($M$ = 2.22, $SD$ = 1.25, $F$ = 12.41, $p$ = 0.006, $\eta^2$ = 0.38). Likewise, conversations in the computer screen condition ($M$ = 3.91, $SD$ = 0.62, $p < 0.001$) were also significantly deeper than those in the poster condition. No significant difference was observed between the robot and the computer screen condition ($p$ = 0.299).

Next, exploratory analysis on the transcripts of participants' conversations using the LIWC software provided additional insights on their conversation experience. In particular, we observed significant differences across experimental conditions in the frequency of emotional tone ($F$ = 3.49, $p$ = 0.040, $\eta^2$ = 0.14), negation ($F$ = 5.05, $p$ = 0.011, $\eta^2$ = 0.20), as well as biological processes ($F$ = 9.69, $p < 0.001$, $\eta^2$ = 0.32) and corresponding subcategories such as body ($F$ = 4.02, $p$ = 0.025, $\eta^2$ = 0.16) and health ($F$ = 8.76, $p < 0.001$, $\eta^2$ = 0.30). Tukey HSD pairwise comparisons further revealed that conversations in the robot condition were significantly more positive in its emotional tone ($p$ = 0.032) and used negation significantly less often than ($p$ = 0.037) those in the poster condition. In addition, participants in the robot condition discussed topics related to biological processes significantly more often than those in the poster condition ($p < 0.001$) or the computer screen condition ($p$ = 0.007). Finally, conversations in the computer screen condition also involved less negation than those in the poster condition ($p$ = 0.016), but differences in other categories did not reach statistical significance. We have also included the results for all LIWC categories as a supplementary document. While the meaning of specific LIWC categories may be open to interpretation and may not completely capture the nuances of the conversations on its own, the patterns that this analysis has revealed suggest that participants in conversations facilitated by the humanoid robot likely had a more positive tone and more frequently discussed topics such as health and body, which were of significant personal relevance during the COVID-19 pandemic. Together with our *Deep Question Compliance Percentage* and human coded depth ratings, these results provide additional support for $\mathbf{H}_3$, $\mathbf{H}_4$, and $\mathbf{H}_5$.

### 4.7 Openness to Further Conversations and Perceived Facilitator Usefulness

In addition to our main conversation experience metrics, participants also reported how their interest in engaging in a conversation with strangers had changed after the study on a scale that ranged from −5 (*decreased greatly*) to +5 (*increased greatly*). Overall, participants indicated that they were more interested in engaging in such experiences. Specifically, we observed a significantly greater increase in the robot condition ($M$ = +2.21, $SD$ = 1.64) than the poster condition ($M$ = +1.13, $SD$ = 1.48, $\beta$ = +1.08, $SE$ = 0.40, $p$ = 0.009). While the robot condition also showed a greater increase than the computer screen condition, this difference was only marginally significant ($M$ = +1.53, $SD$ = 1.50, $\beta$ = −0.68, $SE$ = 0.40, $p$ = 0.096).

We also probed the participants for how useful they thought the conversation facilitators would be in three different scenarios: a networking session, a grocery store, and at home with family members. For a networking session, participants rated the computer screen ($M$ = 7.60, $SD$ = 2.19) to be significantly more useful than the poster ($M$ = 5.87, $SD$ = 3.08, $\beta$ = +1.73, $SE$ = 0.72, $p$ = 0.021). No difference was observed between the robot and the computer screen nor the robot and the poster. In a grocery store, participants believed that the robot ($M$ = 6.46, $SD$ = 2.47) would be significantly more useful than the poster ($M$ = 4.97, $SD$ = 2.87, $\beta$ = +1.50, $SE$ = 0.74, $p$ = 0.049). No difference was observed between the robot and the computer screen nor the computer screen and the poster. For home use, we did not see any significant differences across all three conditions. To summarize, participants rated the computer screen more useful than the poster in networking sessions, and the robot more useful than the poster in a grocery store.

### 4.8 Extraversion and Openness to New Experiences

We also examined how participants' *extraversion* and *openness to new experiences* may have influenced their conversation experience and self-reported interpersonal closeness by introducing these personality traits as additional fixed effects in our statistical analyses. Overall, these personality traits did not influence participants' experience nor perceptions of the robot or their conversation partner except for two exploratory measures. First, those high in extraversion were less likely to mention that they answered the conversations questions when responding to the open-ended question, "Please briefly describe what happened during the time you just spent in the waiting room" ($\beta$ = −13.67, $SE$ = 6.19, $p$ = 0.030). Second, those who scored higher in openness to new experiences were less likely to see the robot as a facilitator ($\beta$ = −16.44, $SE$ = 5.00, $p$ = 0.001) when answering the open-ended question, "How would you best describe the role of the [robot / computer monitor / poster] in the conversation?" ($\beta = -16.44$, $SE = 5.00$, $p = 0.001$).

Most importantly, our findings held true after controlling for those personality traits, which is consistent with research suggesting that both introverts and extroverts can benefit from being prompted to connect with people around them and "act extroverted" [62].

## 5 DISCUSSION

In this study, we examined whether various conversation facilitators – a humanoid robot, a computer screen, and a poster – hold the promise of helping people engage in deeper and more meaningful conversations with strangers in their physical proximity. Given the psychological barriers that frequently keep unacquainted people from engaging in meaningful conversations, such as miscalibrated beliefs that conversations with strangers will go poorly [12, 42] and preconceived notions that strangers are not interested in one's personal revelations [26], we focused our efforts on examining different conversation facilitators' effectiveness at reducing these barriers. Our results demonstrate that technology-enabled conversation facilitators (humanoid robot and computer screen) can elicit greater engagement in conversations about deep and meaningful personal questions, increase interpersonal closeness, and foster a more positive conversation experience between strangers compared with a poster posing the same questions.

Of the three conversation facilitators, the robot facilitator was more effective at getting participants to discuss the two deeper conversation questions than the poster, and it was also marginally more effective than the computer screen, supporting H3. This result is consistent with prior work demonstrating that people are more likely to follow the instructions of a physically embodied robot than a video-based agent [5]. While deploying video based devices (e.g., computer screen) or non-digital devices (e.g., poster) can be simpler and come at a lower cost and on a larger scale, our findings suggest that the humanoid robot could present a unique advantage in facilitating social interactions, especially when such interactions may seem uncomfortable initially (e.g., talking with strangers). Furthermore, although not directly tested in our study, physically embodied agents in real-world deployments can leverage their unique capabilities to apply non-verbal cues — such as gaze, gestures, orientation, or physical proximity — to be even more persuasive and elicit greater compliance from people [9, 14].

Furthermore, accordingly to qualitative coding from three independent coders, participants spent a greater percentage of their conversations discussing personal topics, which suggests that technology-enabled conversation facilitators are better at encouraging self-disclosure between strangers. Consistent with this result, quantitative analysis on participants' conversation transcripts with the Linguistic Inquiry and Word Count (LIWC) software [38] also finds that participants in the robot condition had more positive conversations and more frequently discussed topics of personal relevance (e.g., biological processes). Taken together, these results provide converging evidence that technology has the potential to create opportunities for strangers to connect with each other over deep and meaningful conversations. As such, researchers and public space planners should weigh the unique advantages of technological devices against the lower cost of non-digital devices (e.g., posters) when deciding how to best facilitate meaningful connections among people in public spaces.

Prior work has established that starting conversations with strangers is often difficult and awkward. For example, only 7% of people reported that they are open to talking to strangers in a waiting room [12]. While we expected conversation facilitators with greater social presence (i.e., humanoid robot, computer screen, poster, in that order) to be more successful at getting participants to talk to each other, to our surprise, every participant pair in our study conversed with each other regardless of their experimental condition, which resulted a "ceiling effect" that prevented us from conducting meaningful analyses on H1. Furthermore, all participants reported a similarly high level of easiness and a similarly low level of awkwardness, discomfort, and vulnerability when starting the conversation regardless of the type of conversation facilitator, which led us to reject H2. In retrospect, our small and isolated waiting room evidently made starting conversation with strangers relatively easy—perhaps even somewhat normative. Furthermore, participants were already in a curious and exploratory mindset when joining our study, and as the only two people "stuck" in that room, participants might have felt the need to acknowledge one anothers' presence and strike up a conversation. Therefore, in order to effectively test H1 and H2, future research needs to identify field sites or stage lab environments that more closely resemble a dynamic real-world environment where people would likely experience much greater psychological barriers to engaging in conversations with strangers despite being in the same room.

When it comes to interpersonal closeness (H4) and conversation experience (H5), our results showed that technology-enabled conversation facilitators (humanoid robot and computer screen) not only led to greater reported interpersonal closeness among participants, but also made the conversation experience more positive. This advantage is particularly critical, as various studies have demonstrated the benefits that deep social connectedness can have on one's well-being [3, 11, 25, 36]. Therefore, we suggest that interventions designed to foster communication between strangers in public spaces use interactive displays or technological devices to most effectively promote social connection and foster emotional well-being.

The difference in interpersonal closeness discussed in the previous paragraph was identified using the *Interactive IOS Scale for Multiparty Interactions*, a measure that we developed for this study. It allowed participants to rearrange three circles representing themselves, their conversation partner, and the conversation facilitator to best portray the social dynamics and interpersonal closeness of the different actors in the group. We have made our *Interactive IOS Scale for Multiparty Interactions* measure open source so that others can use and adapt it to measure the interpersonal closeness between interactive agents.

Consistent with previous research on the emotional benefits of talking to strangers, we observed that participants in the robot and computer screen conditions reported feeling significantly more positive after the conversation than before [12]. Curiously, participants in the robot condition, on average, reported feeling both the happiest about the conversation and the most vulnerable during the conversation. While vulnerability is typically considered a negative emotion, it has also been demonstrated to have positive consequences, as it facilitates social engagement and encourages a greater degree of self-disclosure, which may then lead to greater trust within the conversation pair [52, 59]. Possible factors that could contribute to this greater reported happiness are numerous and nuanced and require future research. It is possible that greater compliance with the deeper conversation questions encouraged more self-disclosure, which is an inherently vulnerable undertaking that can contribute to greater happiness and emotional well-being.

Our findings have important implications for designing better social engagement in public spaces. People often spend a substantial portion of their daily life around other people. Despite frequently crossing paths with other individuals in public spaces, people routinely ignore strangers in their close proximity [11, 12, 42], even though studies have shown that engaging in social interactions with other people, strangers included, can have far-reaching physical and mental health benefits [32, 40, 45]. Using a poster as the baseline, we found that humanoid robots and computer screens can bring about greater engagement in deep conversations, increase interpersonal closeness, and foster a more positive conversation experience between strangers. These findings suggest that public space planners should consider leveraging the capabilities of existing and new technologies to encourage social connection and promote emotional well-being. For example, deploying a digital display in a city tour bus that provides conversation prompts and encourages social connection could possibly be enough stimulus to overcome the psychological barriers preventing unacquainted visitors from engaging in meaningful conversations with each other. When planning any real-world deployment, it is also important to consider that different conversation facilitators have unique attributes that resonate better with certain scenarios and spaces than others. As reported by our participants, computer screens are expected to be more useful in networking session scenarios; by contrast, humanoid robots are expected to be more useful in settings like a grocery store. Leveraged effectively, we believe that technological devices can help promote deep social connections and enjoyable conversations between strangers.

## 6 LIMITATIONS

Our study was conducted in a behavioral science center where members of the public could visit on a walk-in basis. It is likely that participants who volunteered to take part in a novel study were already in a relatively exploratory mindset and were thus more inclined to talk to strangers. While we designed the study environment to mimic a real life waiting room, it was likely still evident that our setup was a part of a novel experiment that differed considerably from real-world scenarios (e.g., clinic waiting rooms, coffee shops, commuter trains), where people hold very different expectations about having conversations in public settings [12]. These factors could have contributed to why 100% of participants in our study ended up talking to their conversation partner when only 7% of survey respondents in a prior study reported that they would be open to talking to strangers in a waiting room [12]. As we wanted a controlled experimental environment to bring out any effects the different conversation facilitators may have, conducting the study in a real-world waiting room would have made it impossible to maintain the same level of experimental control. Despite the disparity between our experimental setup and real-world waiting rooms, our findings still hold as our setup was consistent across all three conditions. Furthermore, we anticipate that the humanoid robot will be even more effective in facilitating meaningful conversations between strangers than a computer monitor in a field environment due to the robot's greater social presence. Future work should investigate the efficacy of technology-enabled conversation facilitators in real-world environments.

Additionally, in order to keep the behaviors of all three facilitators closely comparable, we did not leverage the humanoid robot's full range of non-verbal capabilities in this study. Thus, our experiment is a conservative test on the difference between the robot condition and the other two conditions, and differences across conditions can be directly attributed to the delivery medium itself and not their additional features. Even with these conservative design decisions, our results still point to the robot and computer screen being more capable of increasing interpersonal closeness and fostering a more positive conversation experience between strangers. We predict that a humanoid robot with more social interaction capabilities (e.g., gaze, backchanneling, etc.) will perform even better than the two other conversation facilitators. Future research should investigate whether incorporating more social capabilities would enhance the humanoid robot facilitator's effectiveness in helping strangers break the ice. Future research can also examine whether displaying a virtual robot on a computer screen may reap similar benefits as presenting an embodied robot in front of participants.

## 7 CONCLUSION

How technology shapes our social life is a complex and fascinating research question. Some scholars (e.g., Sherry Turkle) have expressed concerns about how "always-on digital connections" can lead us to be "alone together," where people, despite being in the same physical space, are fixated on their own devices and feel miles apart from each other [55]. Although detrimental to people's social and emotional well-being, such tendencies are quite understandable because talking to people who we do not know is difficult, and engaging in *deep conversations* with people is even more daunting.

By contrast, our research shows that technology has the potential to effect positive change by breaking through the social norm of keeping to ourselves that stems from our tendency to overestimate the risks and underestimate the benefits of face-to-face conversations. Specifically, we have investigated whether technologies can "break the ice" between strangers in the same space and help them connect over deep and meaningful conversations. We introduced three conversation facilitators that posed an identical set of progressively deeper conversation questions to pairs of strangers in a waiting room setting. Compared to a poster, a humanoid robot conversation facilitator elicited greater compliance with discussing the deeper questions from the participants. Moreover, participants in conversations facilitated by technology-enabled facilitators such as a humanoid robot and a computer screen also found their conversations to be more enjoyable while also feeling closer to each other and generally happier afterwards, despite feeling more vulnerable with each other during the conversation. Even though technology has been frequently labeled as a major contributing factor to people's declining engagement with each other in close physical proximity, our research demonstrates that technology can also be used as a tool to reduce psychological barriers and to create opportunities for strangers to connect with each other over face-to-face, heart-to-heart conversations.

## ACKNOWLEDGMENTS

We thank the staff at Mindworks for assisting us in conducting experiments for this study. We also thank Nicholas Epley and Michael Kardas for giving us feedback on the research. Finally, we would also like to show our gratitude to Javier Portet, Kaushal Addanki, Liuhao Wu, Bhakti Shah, Stephanie Kim, Efraim Dahl, and Lauren Wright for their assistance in pilot testing, running the study, and analyzing data throughout the study.

## REFERENCES

[1] Monica Anderson, Emily A Vogels, and Erica Turner. 2020. The virtues and downsides of online dating. (2020).
[2] A. Aron, E. N. Aron, and D Smollan. 1992. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. , 596–612 pages. https://doi.org/10.1037/0022-3514.63.4.596
[3] Arthur Aron, Edward Melinat, Elaine N Aron, Robert Darrin Vallone, and Renee J Bator. 1997. The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and social psychology bulletin* 23, 4 (1997), 363–377.
[4] Stav Atir, Kristina A. Wald, and Nicholas Epley. 2022. Talking with strangers is surprisingly informative. *Proceedings of the National Academy of Sciences* 119, 34 (2022), e2206992119. https://doi.org/10.1073/pnas.2206992119 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.2206992119
[5] Wilma Bainbridge, Justin Hart, Elizabeth Kim, and Brian Scassellati. 2011. The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. *I. J. Social Robotics* 3 (10 2011), 41–52. https://doi.org/10.1007/s12369-010-0082-7
[6] Kathy R. Berenson, Anett Gyurak, Özlem Ayduk, Geraldine Downey, Matthew J. Garner, Karin Mogg, Brendan P. Bradley, and Daniel S. Pine. 2009. Rejection sensitivity and disruption of attention by social threat cues. *Journal of Research in Personality* 43, 6 (2009), 1064–1072. https://doi.org/10.1016/j.jrp.2009.07.007
[7] Michael Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Greg Vargas. 2021. 4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community. *Proceedings of the International AAAI Conference on Web and Social Media* 5, 1 (Aug. 2021), 50–57. https://ojs.aaai.org/index.php/ICWSM/article/view/14134
[8] Colleen M. Carpinella, Alisa B. Wyman, Michael A. Perez, and Steven J. Stroessner. 2017. The Robotic Social Attributes Scale (RoSAS): Development and Validation. In *2017 12th ACM/IEEE Int. Conf. on Human-Robot Interaction*. 254–262.
[9] Vijay Chidambaram, Yueh-Hsuan Chiang, and Bilge Mutlu. 2012. Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. 293–300.
[10] Michael Draper. 2008. Self-Disclosure and Friendship Closeness Michael Draper, Rachel Pittard, and Michael Sterling Hanover College 18 April 2008. (2008).
[11] Nicholas Epley, Michael Kardas, Xuan Zhao, Stav Atir, and Juliana Schroeder. 2022. Undersociality: miscalibrated social cognition can inhibit social connection. *Trends in Cognitive Sciences* (2022).
[12] Nicholas Epley and Juliana Schroeder. 2014. Mistakenly seeking solitude. *Journal of Experimental Psychology: General* 143, 5 (2014), 1980.
[13] Miguel Faria, Rui Silva, Patrícia Alves-Oliveira, Francisco S Melo, and Ana Paiva. 2017. "Me and you together" movement impact in multi-user collaboration tasks. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2793–2798.
[14] Francis J Flynn and Vanessa KB Lake. 2008. If you need help, just ask: Underestimating compliance with direct requests for help. *Journal of personality and social psychology* 95, 1 (2008), 128.
[15] Sarah Gillet, Ronald Cumbal, André Pereira, José Lopes, Olov Engwall, and Iolanda Leite. 2021. Robot gaze can mediate participation imbalance in groups with different skill levels. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 303–311.
[16] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality* 37, 6 (2003), 504–528.
[17] Gul Gunaydin, Hazal Oztekin, Deniz Hazal Karabulut, and Selin Salman-Engin. 2021. Minimal social interactions with strangers predict greater subjective well-being. *Journal of Happiness Studies* 22, 4 (2021), 1839–1853.
[18] Chad Harms and Frank Biocca. 2004. Internal consistency and reliability of the networked minds measure of social presence. In *Seventh annual international workshop: Presence*, Vol. 2004. Universidad Politecnica de Valencia Valencia, Spain.
[19] Mitchell Hobbs, Stephen Owen, and Livia Gerber. 2017. Liquid love? Dating apps, sex, relationships and the digital transformation of intimacy. *Journal of Sociology* 53, 2 (2017), 271–284. https://doi.org/10.1177/1440783316662718 arXiv:https://doi.org/10.1177/1440783316662718
[20] Rob W Holland, Ute-Regina Roeder, Baaren Rick B. van, Aafje C Brandt, and Bettina Hannover. 2004. Don't stand so close to me: The effects of self-construal on interpersonal closeness. *Psychological science* 15, 4 (2004), 237–242.
[21] Melissa G Hunt, Rachel Marx, Courtney Lipson, and Jordyn Young. 2018. No more FOMO: Limiting social media decreases loneliness and depression. *Journal of Social and Clinical Psychology* 37, 10 (2018), 751–768.
[22] Daniel Jones. 2015. The 36 questions that lead to Love. https://www.nytimes.com/2015/01/09/style/no-37-big-wedding-or-small.html
[23] Malte Jung, Nikolas Martelaro, and Pamela Hinds. 2015. Using Robots to Moderate Team Conflict: The Case of Repairing Violations, Vol. 2015. https://doi.org/10.1145/2696454.2696460
[24] Younbo Jung and Kwan Min Lee. 2004. Effects of physical embodiment on social presence of social robots. *Proceedings of PRESENCE* 2004 (2004), 80–87.
[25] Michael Kardas, Amit Kumar, and Nicholas Epley. [n. d.]. Digging Deeper: Meaningful Conversations Are Surprisingly Pleasant. ([n. d.]).
[26] Michael Kardas, Amit Kumar, and Nicholas Epley. 2021. Overly shallow?: Miscalibrated expectations create a barrier to deeper conversation. *Journal of Personality and Social Psychology* (2021).
[27] Yutaka Kondo, Kentaro Takemura, Jun Takamatsu, and Tsukasa Ogasawara. 2013. A gesture-centric android system for multi-party human-robot interaction. *Journal of Human-Robot Interaction* 2, 1 (2013), 133–151.
[28] Lara Kroencke, Gabriella M Harari, Mitja D Back, and Jenny Wagner. 2022. Well-being in social interactions: Examining personality-situation dynamics in face-to-face and computer-mediated communication. *Journal of Personality and Social Psychology* (2022).
[29] Ethan Kross, Philippe Verduyn, Gal Sheppes, Cory K Costello, John Jonides, and Oscar Ybarra. 2021. Social media and well-being: Pitfalls, progress, and next steps. *Trends in Cognitive Sciences* 25, 1 (2021), 55–66.
[30] Mark R. Leary. 1983. A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin* 9, 3 (1983), 371–375. https://doi.org/10.1177/0146167283093007
[31] Dong Liu, Roy F Baumeister, Chia-chen Yang, and Baijing Hu. 2019. Digital communication media use and psychological well-being: A meta-analysis. *Journal of Computer-Mediated Communication* 24, 5 (2019), 259–273.
[32] Larisa McLouglin, Barbara Spears, and Carmel Taddeo. 2018. The importance of social connection for cybervictims: how connectedness and technology could promote mental health and wellbeing in young people. *International Journal of Emotional Education* 10, 1 (2018), 5–24.
[33] Matthias R Mehl, Simine Vazire, Shannon E Holleran, and C Shelby Clark. 2010. Eavesdropping on happiness: Well-being is related to having less small talk and more substantive conversations. *Psychological science* 21, 4 (2010), 539–541.
[34] Bilge Mutlu, Toshiyuki Shiwa, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. 61–68.
[35] Tien T. Nguyen, Duyen T. Nguyen, Shamsi T. Iqbal, and Eyal Ofek. 2015. The Known Stranger: Supporting Conversations between Strangers with Personalized Topic Suggestions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 555–564. https://doi.org/10.1145/2702123.2702411
[36] Marta Pancheva and Alejandra Vásquez. 2022. Close to others-closer to happiness?: An empirical investigation of the social determinants of subjective wellbeing. *International Journal of Wellbeing* 12, 2 (2022).
[37] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015.* Technical Report.
[38] James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates* 71, 2001 (2001), 2001.
[39] Anat Perry, David Mankuta, and Simone G Shamay-Tsoory. 2015. OT promotes closer interpersonal distance among highly empathic individuals. *Social cognitive and affective neuroscience* 10, 1 (2015), 3–9.
[40] Sarah D Pressman, Sheldon Cohen, Gregory E Miller, Anita Barkin, Bruce S Rabin, and John J Treanor. 2005. Loneliness, social network size, and immune response to influenza vaccination in college freshmen. , 297-306 pages. https://doi.org/10.1037/0278-6133.24.3.297

[41] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction* 1, 1 (2012), 119–136.
[42] Gillian M Sandstrom and Erica J Boothby. 2021. Why do people avoid talking to strangers? A mini meta-analysis of predicted fears and actual experiences talking to a stranger. *Self and Identity* 20, 1 (2021), 47–71.
[43] Gillian M Sandstrom, Erica J Boothby, and Gus Cooney. 2022. Talking to strangers: A week-long intervention reduces psychological barriers to social connection. *Journal of Experimental Social Psychology* 102 (2022), 104356.
[44] Gillian M Sandstrom and Elizabeth W Dunn. 2014. Is efficiency overrated? Minimal social interactions lead to belonging and positive affect. *Social Psychological and Personality Science* 5, 4 (2014), 437–442.
[45] Gillian M Sandstrom and Elizabeth W Dunn. 2014. Social interactions and well-being: The surprising power of weak ties. *Personality and Social Psychology Bulletin* 40, 7 (2014), 910–922.
[46] Holly Schiffrin, Anna Edelman, Melissa Falkenstern, and Cassandra Stewart. 2010. The associations among computer-mediated communication, relationships, and well-being. *Cyberpsychology, Behavior, and Social Networking* 13, 3 (2010), 299–306.
[47] Juliana Schroeder, Donald Lyons, and Nicholas Epley. 2022. Hello, stranger? Pleasant conversations are preceded by concerns about starting one. *Journal of Experimental Psychology: General* 151, 5 (2022), 1141.
[48] Sarah Sebo, Brett Stoll, Brian Scassellati, and Malte F Jung. 2020. Robots in groups and teams: a literature review. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–36.
[49] Solace Shen, Petr Slovak, and Malte F Jung. 2018. " Stop. I See a Conflict Happening." A Robot Mediator for Young Children's Interpersonal Conflict Resolution. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 69–77.
[50] Donghoon Shin, Sangwon Yoon, Soomin Kim, and Joonhwan Lee. 2021. BlahBlahBot: Facilitating Conversation between Strangers Using a Chatbot with ML-Infused Personalized Topic Suggestion. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 409, 6 pages. https://doi.org/10.1145/3411763.3451771
[51] Sarah Strohkorb Sebo, Ling Liang Dong, Nicholas Chang, and Brian Scassellati. 2020. Strategies for the inclusion of human members within human-robot teams. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 309–317.
[52] Sarah Strohkorb Sebo, Margaret Traeger, Malte Jung, and Brian Scassellati. 2018. The ripple effects of vulnerability: The effects of a robot's vulnerable behavior on trust in human-robot teams. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 178–186.
[53] Jan Svennevig. 1999. *Getting Acquainted in Conversation: A Study of Initial Interactions*. John Benjamins, Amsterdam. 384 pages. https://benjamins.com/catalog/pbns.64
[54] Margaret L Traeger, Sarah Strohkorb Sebo, Malte Jung, Brian Scassellati, and Nicholas A Christakis. 2020. Vulnerable robots positively shape human conversational dynamics in a human–robot team. *Proceedings of the National Academy of Sciences* 117, 12 (2020), 6370–6375.
[55] Sherry Turkle. 2012. *Alone together: why we expect more from technology and less from each other.* Basic Books. 384 pages.
[56] Jean M Twenge, Jonathan Haidt, Andrew B Blake, Cooper McAllister, Hannah Lemon, and Astrid Le Roy. 2021. Worldwide increases in adolescent loneliness. *Journal of adolescence* 93 (2021), 257–269.
[57] Daniel Ullman and Bertram Malle. 2018. What Does it Mean to Trust a Robot?: Steps Toward a Multidimensional Measure of Trust. 263–264. https://doi.org/10.1145/3173386.3176991
[58] Dina Utami and Timothy Bickmore. 2019. Collaborative user responses in multiparty interaction with a couples counselor robot. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 294–303.
[59] Gerben A Van Kleef, Carsten KW De Dreu, and Antony SR Manstead. 2010. An interpersonal approach to emotion in social decision making: The emotions as social information model. In *Advances in experimental social psychology*. Vol. 42. Elsevier, 45–96.
[60] Chia-chen Yang. 2016. Instagram use, loneliness, and social comparison orientation: Interact and browse on social media, but don't compare. *Cyberpsychology, behavior, and social networking* 19, 12 (2016), 703–708.
[61] Zhimin Yang, Boying Zhang, Jiangpeng Dai, Adam C. Champion, Dong Xuan, and Du Li. 2010. E-SmallTalker: A Distributed Mobile System for Social Networking in Physical Proximity. In *2010 IEEE 30th International Conference on Distributed Computing Systems*. 468–477. https://doi.org/10.1109/ICDCS.2010.56
[62] John M Zelenski, Deanna C Whelan, Logan J Nealis, Christina M Besner, Maya S Santoro, and Jessica E Wynn. 2013. Personality and affective forecasting: Trait introverts underpredict the hedonic benefits of acting extraverted. *Journal of Personality and Social Psychology* 104, 6 (2013), 1092.
[63] Xuan Zhao and Nicholas Epley. 2021. Insufficiently complimentary?: Underestimating the positive impact of compliments creates a barrier to expressing them. *Journal of Personality and Social Psychology* 121, 2 (2021), 239.
[64] Xuan Zhao and Nicholas Epley. 2022. Surprisingly happy to have helped: How underestimating prosociality creates a misplaced barrier to asking for help. (2022).