Sharing Stress With a Robot: What Would a Robot Say?

Honson Y. Ling[1] and **Elin A. Björling**[1]

1 Department of Human Centered Design & Engineering, University of Washington, Seattle, Washington, United States of America

Abstract

With the prevalence of mental health problems today, designing human-robot interaction for mental health intervention is not only possible, but critical. The current experiment examined how three types of robot disclosure (emotional, technical, and by-proxy) affect robot perception and human disclosure behavior during a stress-sharing activity. Emotional robot disclosure resulted in the lowest robot perceived safety.

Post-hoc analysis revealed that increased perceived stress predicted reduced human disclosure, user satisfaction, robot likability, and future robot use. Negative attitudes toward robots also predicted reduced intention for future robot use. This work informs on the possible design of robot disclosure, as well as how individual attributes, such as perceived stress, can impact human robot interaction in a mental health context.

Keywords: social robots, self-disclosure, stress, human-robot interaction, teleoperation, attitudes towards robots, robot dialogue design

Introduction

The rapid development of robotics promises a diverse integration of robots into our daily life. The psychological and social benefits associated with interacting and communicating with a sociable machine have fascinated researchers in psychology, human-robot interaction (HRI), and human-machine communication (HMC) (Guzman, 2018; Mitsunaga et al., 2006). Past research has demonstrated social robots' capabilities to further psychological well-being in vulnerable populations. Social robots are not only capable of evoking empathy, they can decrease loneliness in the elderly, improve social capabilities of older people with dementia, elicit novel social behavior from people with autism, and foster social engagement and self-disclosure among adolescents (Chu et al., 2017; Kwak et al., 2013; Martelaro et al., 2016; Robinson et al., 2013; Rose & Björling, 2017; Scassellati et al., 2012). While studies

have shown that sharing stressful experiences can help reduce stress, few studies have delved into the design of a stress-sharing interaction between a human and a robot (Hofmann et al., 2012; J. H. Kahn, Achter, & Shambaugh, 2001; J. H. Kahn & Hessling, 2001; Zhang, 2017). With nearly one in five US adults living with a mental illness (46.6 million in 2017), designing human-robot interaction for mental health intervention (MHI) is not only possible, but critical for societal and individual well-being (NIMH, 2017).

Based on our survey of the existing literature, we discovered two factors that are understudied but critical in designing HRI for a mental health context: the role of individual attributes (who the users are) and the role of robot disclosure (what the robot says). In this paper, we describe our design and study of a stress-disclosure interaction between a human and a robot. We investigated the following questions: (1) How do different types of robot self-disclosure affect human disclosure behavior and their perception of robot disclosure? (2) How do individual attributes (such as shyness, stress level, and attitude toward robots) affect human disclosure behavior and their robot attributes?

In our background section, we surveyed relevant works on self-disclosure in human-human and human-robot interaction, as well as the role of individual attributes in HRI. Then, we present two central research questions on robot disclosure and human individual attributes motivated by the fields of HMC and HRI. In the methodology and analysis section, we present our design of a small pilot study to explore the proposed research questions. Finally, we present our findings and discussion during the results and discussion sections. Ultimately, the aim of this paper is to contribute to the field of human-machine communication by addressing complexities, such as the effect of human stress and human interpretation of robot emotions, in designing a stress-sharing activity between a human and a robot.

Background and Related Work

The topic of self-disclosure and mental health has been widely studied from a clinical and psychological context. Self-disclosure involves the act of revealing personal information about oneself to another agent, typically a human (Collins & Miller, 1994). Such actions have been viewed as central to the development of close relationships and to the maintenance of psychological well-being (Altman & Taylor, 1973; Jourard, 1964). Disclosure can improve an individual's self-image, such as experiencing greater self-affirmation thereby restoring a sense of worth after intimate disclosure (Creswell et al., 2007).

The Benefits of Self-Disclosure

Self-disclosing personal stress is an effective way for people to reduce and manage their stress (Hofmann et al., 2012; J. H. Kahn, Achter, & Shambaugh, 2001; J. H. Kahn & Hessling, 2001; Zhang, 2017). More specifically, the act of disclosing intimate and emotional information is associated with decreased depressive symptoms (J. H. Kahn & Garrison, 2009). The benefits of emotional disclosure also include the improvement of immune function, decrease of emotional and physical symptoms resulting from trauma, and protection against depression (Esterling et al., 1994; Pérez et al., 2017). Furthermore, Esterling et al. have found that verbal expression about stressful events, compared to written expression, achieved greater improvements in cognitive change, self-esteem, and adaptive coping strategies. More recently, talking with an online chatbot has been shown effective in reducing

participants' stress (Fitzpatrick et al., 2017; Huang et al., 2015). Therefore, the current study explores the idea of designing a social robot that engages and encourages users to self-disclose in a stress-sharing activity.

Self-Disclosure in Human-Computer Interaction

In order to encourage people to talk to a robot, we reviewed a body of research and found that self-disclosure can be elicited through reciprocation from a conversation partner (Altman & Taylor, 1973; Collins & Miller, 1994; Taylor & Hinds, 1985). In other words, a person is more likely to self-disclose if the conversational partner also engages in self-disclosure. This reciprocal phenomenon occurs not only in human-human interactions, but also between human and technologically-mediated social agents.

Studies have shown that during a technologically-mediated social interaction, people tend to share more to mediated agents (such as a chatbot, other online forum users, or social robots) that also share about themselves (Barak & Gluck-Ofri, 2007; Martelaro et al., 2016; Moon, 2000). Hence, implementing self-disclosure behavior in a social robot for MHI may encourage users to share in greater length and emotional depth about their personal stress, which may also lead to greater psychological benefits, such as cognitive change, self-esteem, and adaptive coping strategies.

When it comes to the design of robot self-disclosure, there have been few studies on what a robot should self-disclose to people in a stress-sharing context. While the role of a social robot for mental health is far from that of a human therapist, therapist self-disclosures are well-documented (Goldfried et al., 2003; Henretty & Levitt, 2010). From this literature, therapist self-disclosure can vary in terms of intimacy (depth), duration (breadth), timing, content, and so forth (Henretty & Levitt, 2010). Overall, careful therapist self-disclosure can benefit the overall quality and experience of the client. As Goldfried et al. (2003) suggests, "therapist self-disclosure emerges as a natural part of the intimate, human interaction of therapy" (p. 567).

Outside of the realm of therapy, past studies in human-computer interaction have often adopted the Computer Agent as Social Actor (CASA) framework to investigate a user's verbal behavior and interaction with computer agents. This framework proposed that people instinctively perceive, react to, and interact with computers as they do with other people, without consciously intending to do so (Reeves & Nass, 1996).

Specifically, Moon (2000) compared human disclosure in a reciprocal versus non-reciprocal computer condition. In the reciprocal condition, the computer preceded each question with some technical information about itself, such as "This computer has been configured to run at speeds up to 266 MHz." Moon found that a computer which disclosed information about itself resulted in greater depth and breadth of participant responses and higher ratings of likability compared to a computer offering no disclosure. In another study, Ho et al. (2018) randomly assigned participants to interact with a confederate on an online chat platform who was either perceived as a chatbot or a real human actor. They found that the conversation on the platform was effective in creating relational, emotional, and psychological benefits, regardless whether the conversational partner was perceived as a human or a chatbot. Furthermore, they found that the emotional condition (in which the confederate provided participants with validating responses and asked more probing questions) elicited more disclosure, enhanced perceived understanding and disclosure intimacy

between the partners, compared to the factual condition (in which the confederate did not ask about participants' feelings or emotions).

Self-Disclosure in Human-Robot Interaction

Through our literature review on self-disclosure in HRI, we encountered three areas of existing work on this topic: (1) how nonverbal robot behaviors, such as eye gaze, affect human disclosure, (2) how robot disclosure affects different human outcomes, and (3) how human attitudes toward the robot affect human disclosure to the robot.

In the area of nonverbal robot behaviors, researchers have explored how physical distancing, eye gaze, hugs, and physical presence affect human self-disclosure (Mumm & Mutlu, 2011; Pettinati et al., 2016; Powers et al., 2007; Shiomi et al., 2017). Although the findings vary widely, nonverbal robot behavior, such as hugging, have been found to increase human disclosure. While this first area is critical for the design and implementation of a social robot for MHI, existing research related to robot disclosure is scarce. Due to the increasing relevance of linguistic communication between human and social robots, as well as the potential implication of human-machine communication for MHI, we focused our current study on how the content of robot disclosure, as well as human attributes, affect human-robot interaction outcomes (Sandry, 2018).

Among studies looking at the effect of robot self-disclosure, Nomura and Kawakami (2011) found that negative robot disclosure or no robot disclosure increased human anxiety after robot interaction, while positive robot disclosure did not. In the positive self-disclosure condition, the robot uttered its recent positive situation ("I am very fine due to the maintenance conducted a few days ago.") compared to uttering a recent negative situation in the negative disclosure condition ("My motors are not well, but have still not been restored."). While human anxiety is affected, subjects' self-disclosure behaviors toward the robot were not affected by either type of robot disclosure in this short, single-response interaction.

In another study, Mumm and Mutlu (2011) manipulated a robot into either likable (a polite, empathetic 20-second monologue during the introduction) or unlikable (a rude, selfish monologue) behavior and found people answered more sensitive questions from the likeable robot. In a most recent study, Johanson et al. (2019) found that a health care robot using self-disclosure and a forward lean increased human engagement and attentional behaviors. These studies not only provide a glimpse of the possible effect of different robot disclosures on human anxiety toward the robot and disclosure behavior, they also support the CASA framework which suggests that people attribute verbal communication from a social robot to inform their own perception of the robot and their behavior toward it.

Two existing studies investigate even more closely on the topic of robot disclosure by manipulating the intimacy or vulnerability level of the robot's disclosure. In a small exploratory study ($n = 11$), Burger et al. (2016) found that diabetic children were less likely to respond to increased intimacy in robot disclosure. The authors provide several explanations, such as children may have felt overwhelmed by higher intimacy disclosures ("too much information") or that children wanted to match the robot's intimacy but weren't capable of sharing on that level.

In another study with children, Martelaro et al. (2016) found that high school students ($n = 61$) disclosed more about their own vulnerability when interacting with a robot that

discloses high vulnerability rather than one that discloses low vulnerability. They created high robot vulnerability through statements that convey perceived weakness such as, "Every time I run a new program I get a bit stressed," and low robot vulnerability through factual statements such as, "Each new program I run changes what I can do." In this study, robot vulnerability was associated with increased ratings of trust and companionship, suggesting that designing robot vulnerability is a factor of building companionship between humans and robots. These two studies suggest that human attributes, as discussed later, play a critical role in human self-disclosure with robots.

On the topic of designing a series of robot disclosures, Ligthart et al. (2019) proposed five interaction design patterns (IDPs) that focus on the process of getting acquainted between a user and a robot. These IDPs touch upon what questions a robot should ask, what responses it should give, and what structure the conversation should take hold as. More specifically, they suggest that during a getting acquainted interaction, a robot should pair closed-ended and open-ended questions, acknowledge participant responses, and engage in a six-step turn-taking mechanism. The combination of these IDPs provide a structure for the robot to autonomously process self-disclosures from people, while also being stimulating for people to engage in the conversation.

In sum, the above findings suggest that carefully implemented, appropriate forms of robot self-disclosures over a sufficient length of interaction time can successfully elicit human self-disclosure and potentially positive robot attributions. However, a series of questions related to robot disclosure remains. If vulnerability really causes people to disclose more, how do we operationalize vulnerability for a robot? Should a robot disclose emotions like humans do? How would people interpret emotions from a social robot? While the current study cannot possibly provide definitive answers to these questions relating to the ontological classification and ethics of human-machine communications, understanding the effect of different types of robot disclosure might be the first step to unravel these questions. In addition, understanding the design of robot disclosure will undoubtedly provide insights to the implementation of social robots in mental health. In the sensitive context of mental health intervention, designing appropriate verbal communication should be the priority in avoiding causing human harm. Nash et al. (2018) have found that verbal social rejection from a social robot following a game with the robot decreases participants' self-esteem. Thus, this current paper highlights this necessity by exploring the effect of different robot disclosure designs.

Human Attributes in Human-Robot Interactions

Aside from robot behaviors, human individual characteristics, such as personality, stress level, and general attitudes toward robots, also play a critical role in HRI, especially in the context of mental health. In terms of personality, Salem et al. (2015) found that extroverts felt psychologically closer to the robots, compared to introverts, during the robot interactions. Ligthart et al. (2019) found that extroverted children self-disclose more to robots. More specifically, Nomura, Kanda, et al. (2008) found that people with higher negative attitudes and anxiety toward interaction with robots tend to avoid talking with a robot. In combination with personality, these are the human characteristics that have been commonly studied in the HRI literature. On the other hand, perceived stress is a factor that is

less commonly studied within the HRI community. Perceived stress is an important indicator of the degree to which situations in one's life are appraised as stressful (Cohen et al., 1994). High perceived stress is associated with greater vulnerability to stressful life-event-elicited depressive symptoms and health-related issues. To the best of our knowledge, there has been limited exploration of perceived stress and self-disclosure in HRI. While it has been shown that social robots can reduce stress as a result of longer-term interactions in the elderly and promote more physical movement and more emotional verbal expressions in children (Jeong, 2017; Wada et al., 2005), Jeong did not find any significant change in children's perceived stress (but found changes in affect and user engagement) during a three-week longitudinal study of implementing a virtual avatar that employs mental health intervention strategies via verbal interactions. From a mental health perspective, it is critical not only to understand whether interacting with a robot can reduce stress, it is also important to examine how an individual's stress level affects their behaviors and perception of robots during a human-robot interaction.

Research Questions

The current paper presents a pilot study that contributes to the fields of HMC and HRI in two areas, the effect of different types of robot disclosure, and the effect of individual characteristics on human-robot disclosure. More specifically, the study investigated (*RQ1*) How does the type of robot disclosure affect (1a) human disclosure (length and depth) and (1b) perception of the robot? (*RQ2*) How do individual characteristics (shyness, stress level, and attitude toward robots) affect (2a) human disclosure and (2b) perception of the robot?

According to the computers as social actors (CASA) framework, the effects of emotional disclosure should operate in the same way for human and technological social agents (Ho et al., 2018; Kang & Gratch, 2010; Von der Puetten et al., 2010). Thus, emotional robot disclosure would, in theory, elicit increased length and depth of human disclosure, as well as increased likability for the robot, compared to robot disclosure that does not involve emotions. For *RQ1a*, we hypothesized that emotional robot disclosure will elicit the longest and deepest participant disclosure, with technical disclosure eliciting the least depth and breadth of disclosure. For *RQ1b*, we hypothesized that emotional robot disclosure will lead to the highest positive user perception of the robot (such as likability, perceived safety, user satisfaction, and intention for future use), with technical disclosure eliciting the least positive user perception.

For *RQ2*, due to the exploratory nature of this question, we do not have any directional hypotheses, but are merely interested in whether individual attributes (shyness, stress level, and attitude toward robots) have an effect on human disclosure behavior or user perception of the robot.

Designing Robot Interaction and Disclosures

There are two important design components in this study: the stress-sharing human-robot interaction as a whole, and the types of robot disclosure with varying degrees of intimacy. The holistic design rationale of the stress-sharing interaction follows the "getting acquainted between a user and a robot" interaction design patterns, ordering the conversation with greeting and introductory questions, such as "How are you doing?" to increasingly intimate

questions, such as "Have you ever felt overwhelmed" over the course of the conversation (Ligthart et al., 2019; Moon, 2000).

The basic structure of the conversation across all conditions was comprised of: (1) greeting and opening questions, (2) first robot-disclosure statements (designed for each condition), (3) an open-ended question asked by the robot (similar across conditions), (4) participant's response, (5) a generic robot response (e.g., "okay," "I see," "interesting") (similar across conditions). This question-response turn-taking style has been documented by Ligthart et al. (2019). The robot asked a total of four open-ended questions: how the participants feel about his/her age, hometown, hobbies, and stress.

Regarding stress, the robot used two probing techniques (e.g., "Could you tell me more?") to encourage self-disclosure from the participants about their recent stressful experience. In order to allow for a natural interaction, the operator waited approximately 3 seconds before each response to ensure the participant was done talking. See Appendix for the complete robot script.

Given the novelty of designing for different types of robot disclosure, we drew from limited resources for guidelines and best practices from HCI and previous social psychology research on factual and emotional disclosure (Barak & Gluck-Ofri, 2007; Laurenceau et al., 1998). We translated existing human disclosure categories into context-appropriate disclosure categories for a collocated, social robot. We created a third classification of "by-proxy" disclosure. In this novel disclosure category, the robot shares another person's data, via a stressful experience as a form of disclosure. In other words, the robot is detached from mentioning its own emotional state, but rather acts as a medium to relate users through other users' feeling. We chose to include this new type of disclosure as it might be contextually appropriate for social robots as a medium, rather than an entity. In addition, future personalized robots or virtual assistants might be designed with functions to inform and relate users about others' emotional states similar to the current use of social media (Stieglitz & Dang-Xuan, 2013). See Table 1 for details. We have also maintained the word count of each robot disclosure to be similar across conditions. These three types of robot disclosure were successfully identified by a small set of people ($n = 3$) that were not involved in the experimental design process.

TABLE 1   Description of Robot Disclosure

Type of Robot Disclosure	Description	Dialogue Example
Emotional Disclosure	Robot shares its own relatable experience and feelings.	*"Recently, I had to juggle between multiple programs at once through my system. I was quite overwhelmed because I felt like I had too much on my plate."*
By-Proxy Emotional Disclosure	Robot shares other users' relatable experience from prior encounter.	*"Recently, I have talked to people that had to juggle many things in life. They were quite overwhelmed because they felt like they had too much on their plate."*
Technical Disclosure	Robot shares information about its technical specification, functions, or past events.	*"Recently, I had to juggle between multiple programs at once. My system was overwhelmed and crashed because my battery became overheated. I was unable to function properly."*
Robot word count is controlled across conditions (+/− 1word)		

Robot Specification

In this study, we utilized an existing robot prototype, named EMAR V4, which was originally designed and developed to gather stress data from teens (Rose & Björling, 2017). EMAR V4 is a social robot designed for ease of programming and customization. It has two Nexus 7 tablets cased in a soft felt body. One tablet is used as the robot's face, which is a web application running on a browser on the tablet. Features of the face can be modified through a browser-based interface that communicates with the face tablet through a real-time database. The face has two eyes that blink and its facial expression can be changed. The face tablet is also used to project the voice of the robot using the browser's text-to-speech capability. During this experiment, both the facial expression and the androgynous voice remained consistent for all participants.

The other tablet is located at the robot's belly and is intended as an input/output touch-screen for communication with the user. In this study, the belly tablet was used to display what the robot said in text form, similar to subtitles. The robot's responses are controlled by the experimenter through another browser-based interface. For our study, the interface was populated with the pre-specified responses that the experimenter could choose from in each condition in response to the participant's utterance, to enable a fluent interaction. Nonetheless, the interface included a free-form text box response to address unexpected participant questions. See Figure 1 for visual detail.

Participants

A total of 36 participants (52.8% women, *M* age = 21.6) were recruited from a university through convenience sampling using emails, flyers, and word-of-mouth during the summer of 2018. The self-described ethnicities of our sample consisted of 67% Asian, 23% White, 5% Black, and 5% other. The study was approved by the university's institutional review board. Participants gave verbal consent before the researcher began the study. Participants were compensated with a $10 gift card at the end of the study.

FIGURE 1   Left: laptop displaying the control interface used in the experiment; right: how the robot V4 looks during the experiment.

Study Procedures

Each participant was randomly assigned to one of the three experimental conditions: emotional, by-proxy, or technical disclosure. Condition groupings did not significantly differ in terms of gender [$\chi^2(2, N = 36) = 1.56, p = .458$] or age [$F (2, 33) = .914, p = .411$]. Upon arriving at the research building, each participant was greeted by an interviewer and led to a room to read a consent form informing him or her about the experimental task, which involved interacting with a social robot prototype capable of engaging in a conversation with people. Participants were then asked to complete a computer-based intake-questionnaire that captured their demographic information, shyness, perceived stress level, and negative attitudes toward robots.

After the participant completed the initial surveys, the interviewer introduced EMAR V4 and exited the room. Meanwhile another researcher began the wizard-of-oz (Dahlbäck et al., 1993) control of the robot through a web-based interface controller with a live audio feed in the room. See Figure 2 for the experimental room layout.

The trained robot operator followed a specified script (see Appendix) designed based on the rationale mentioned previously. Upon completion of the interaction stage, the researcher re-entered the room and asked the participant to fill out the post-interaction questionnaires. Participants were then asked four open-ended questions about their experience. Participants were fully debriefed, and were told that the robot was controlled by the experimenter through a script. We then answered any questions they had for the study before concluding the experiment.

Intake Instruments.   In order to capture potential moderators and the variables of interest, our intake survey captured basic demographic information on participant's age, gender, and ethnicity, as well as individual characteristics, such as shyness, stress level, and attitude toward robots before participants interacted with the robot. Participants' self-reported stress level was captured using the Perceived Stress Scale Cohen et al. (1994), a 10-item questionnaire that measures the degree to which situations in one's life are appraised as
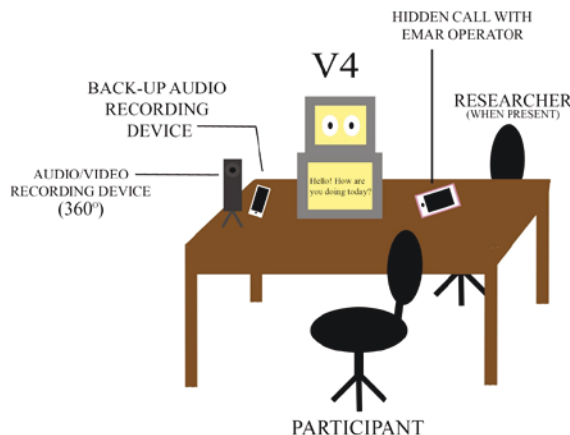
FIGURE 2   Experimental room layout diagram. The robot's head height is roughly in line with the seated human's eyes.

stressful. Participants' shyness was scored using the 13-item revised Cheek and Buss shyness scale (RCBS) from Cheek (1983). Finally, participants completed the 14-item Negative Attitudes Toward Robots Scale (NARS) from Nomura, Suzuki, et al. (2006).

Behavioral Measures During Interaction.   To capture each participant's level of disclosure, we video recorded and transcribed each human-robot interaction. The analysis of participant disclosure behavior will be explained in the next section.

Post-Interaction Instruments.   In order to investigate *RQ1*, participants completed a survey after the robot interaction on user satisfaction (four items, like "I feel absorbed in the conversation with [V4]") and intention for future use (four items, like "I will use [V4] again") adapted from Lee and Choi (2017). We decided to utilize the user satisfaction scale because Lee and Choi have found that reciprocity and self-disclosure are strong predictors of relationship building and user satisfaction between users and virtual agents. Therefore, user satisfaction might be an important indicator toward understanding the nature of human-robot interaction involving the reciprocity of disclosure. While Lee and Choi utilized the intention for future use scale to measure acceptance and potential loyalty to a virtual assistant agent, this construct is important for understanding how users feel about interacting with a robot again in the future. To understand more about the perception of robot attributes, we decided to measure robot likability (five items) and robot perceived safety (two items) taken from Bartneck et al. (2009). The likability scale asks participants to rate their impression of the robot with items such as dislike/like, unkind/kind, and so forth. The perceived safety scale measures the user's perception of the level of danger when interacting with a robot, and the user's level of comfort during the interaction (Bartneck et al., 2009). It asked participants to rate their affective state with items such as anxious/relaxed, calm/agitated. Overall, the above scales inform not only the quality of the interaction, but also participant's perception of the robot's attributes.

Manipulation Check.   In order to make sure that conditions were successfully manipulated, one question in the post-interaction survey asked participants, "Which of the following best describes EMAR V4's style of communication?" with options: robot tends to talk about "its own emotion," "the experience of others," or "technical information about its system and programs."

Brief Interview.   Given the novelty of this study, we included a brief exit interview after the post-interaction survey to get a qualitative understanding of how participants felt about the interaction. The researcher asked four open-ended interview questions after the interaction to capture participant's interaction experience (Birnbaum et al., 2016). The interview questions were: "How was your experience with the robot?," "How did it make you feel to talk and disclose about yourself to the robot?," "How did you perceive the robot's personality?," "Would you disclose to robots in the future? Why or why not?" These interviews were recorded and transcribed for analysis.

Analyses

Data Management

All survey data were coded, scored if appropriate (PSS, NARS, RCBS) using R (version 3.6.0). Data were then cleaned and explored for outliers and normalization. Analysis was done in R and later cross-referenced by another researcher using SPSS version 24 to ensure the accurate results.

Participant Disclosure Analysis

Transcript of participants' responses to the robot were analyzed for the degree of self-disclosure through the dimensions of length and depth according to previous self-disclosure studies (Collins & Miller, 1994; Ho et al., 2018). Length refers to the quantity of the information exchanged and is often measured using a word count, whereas depth refers to the quality of the information disclosed and is often measured using an intimacy scheme (Altman & Taylor, 1973; Collins & Miller, 1994; Kang & Gratch, 2010).

To obtain disclosure length, we counted the total amount of words spoken to the robot by each participant. To obtain disclosure depth, we first coded our transcripts into utterances, which is defined as one complete sentence or phrase (Guetzkow, 1950). A second coder coded 20% of the transcripts into utterances. This yielded a Cohen's kappa of .75. As a result, a total of 3,259 utterances were coded. The average length of an utterance was 13 words and, on average, participants spoke 33.26 utterances. Then, we rated each disclosure utterance into three levels of intimacy: low level, which includes objective facts about the situation; medium level, which includes attitudes, thoughts, and opinions about the situation; and high, which consists of explicitly verbalized emotions and affect. Utterances that were not disclosure (e.g., "thank you") were coded as 0. A second coder coded 20% of the utterances, yielding a Cohen's kappa of .85. Each utterance received a score from 0 to 3. Scores for disclosure-only statements (1–3) were averaged and then normalized according to the number of disclosure utterances each participant gave in the conversation, such that each participant received an overall disclosure depth score.

Exploring Group Differences Through Quantitative and Qualitative Analysis

After coding and scoring all of the raw data, we conducted a Spearman rank correlation analysis in order to explore how individual characteristics such as shyness, perceived stress, and negative attitudes toward robots (NARS) were related to interaction and outcome measures. Then, a GLM multivariate test was conducted to test if word count, disclosure depth, user satisfaction, intention for future use, likability, and perceived safety differed based on robot disclosure conditions, with the covariates of perceived stress, robot attitudes (NARS), and shyness. Lastly, a collaborative applied thematic analysis was used to explore the qualitative data (e.g., conversations with the robot and interview responses) to further explore the nature of human robot interaction in the context of a stress intervention (Guest et al., 2011).

Results

All participants appeared comfortable and engaged in their conversations with the robot and all responded to each question asked by the robot. The length of robot interactions ranged between 2:21 and 9:54 minutes ($M = 4.20$) and participant word count ranged from 31 to 1,029 words ($M = 219$).

Correlations

During our correlation analysis, we discovered several statistically significant correlations between individual participant characteristics and experimental outcomes. Shyness was negatively correlated with disclosure depth and perceived safety. NARS was negatively correlated with intention for future use. Surprisingly, perceived stress was positively correlated with shyness, and negatively correlated with shyness, disclosure length, disclosure depth, user satisfaction, likability, and future use. See the correlation matrix (Table 2) for details.

Robot Disclosure Condition Confusion

The potential control of our manipulation needs to be interpreted with care as 44% ($n = 16$) of our participants failed to correctly identify the type of robot disclosure they experienced. Participants' ability to correctly identify the type of robot disclosure was not statistically different across the three robot conditions [$\chi^2(2, N = 36) = 4.275, p = .118$]. Although 9 of the 12 (75%) participants correctly identified the type of robot disclosure in the by-proxy condition, only 4 out of 12 participants (33%) in the emotional condition and 7 out of 12 (58%) in the technical condition correctly identified the type of robot disclosure in their corresponding assigned conditions. More specifically, 4 out of 12 participants interpreted the technical robot disclosure as emotional robot disclosure, while 5 out of 12 interpreted emotional robot disclosure as technical. See Figure 3 for a confusion matrix with full details. We offer several plausible explanations for this phenomenon in our discussion section.

TABLE 2   Correlation Matrix

	PSS	NARS	Shyness	Disclosure Length	Disclosure Depth	User Satisfaction	Likability	Future Use
PSS								
NARS	0.20							
Shyness	0.61***	0.35*						
Disclosure Length	-0.45**	0.12	-0.29					
Disclosure Depth	-0.48**	0.09	-0.35*	0.94***				
User Satisfaction	-0.50**	-0.32	-0.23	0.16	0.14			
Likability	-0.44**	-0.11	-0.28	0.26	0.24	0.70***		
Future Use	-0.50**	-.39*	-0.19	-.14	0.13	0.75***	0.67***	
Safety	-0.19	-0.29	-0.40*	-.03	0.20	-0.03	-0.06	0.03
*Note: *p<0.05; **p<0.01; **p<0.001*								
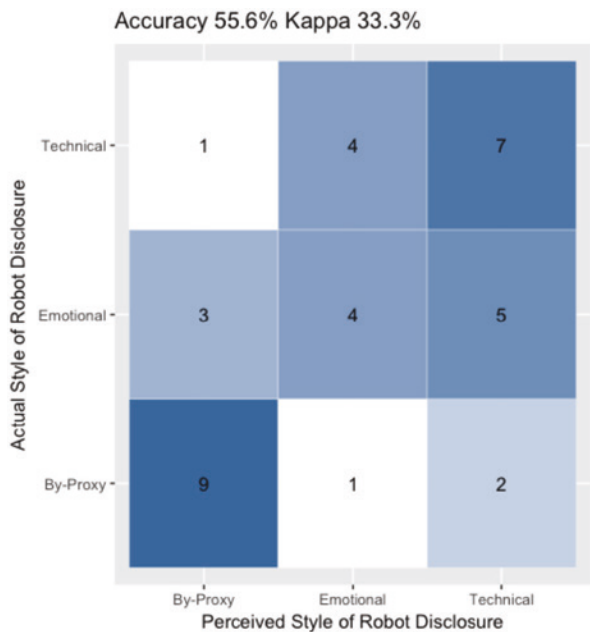
FIGURE 3   Confusion Matrix on the actual style of robot disclosure compared to the perceived style of robot disclosure by the participants.

Effect of Robot Disclosure Condition

Although the condition confusion needs to be taken into consideration, a multivariate test of robot perceptions revealed that robot perceived safety was significantly higher for participants who had interacted with the robot in the technical condition [$F(2) = 3.684, p = .037$] followed by the by-proxy condition and the emotional condition. No significant differences in the length of human disclosure (word count) or depth of disclosure were found across conditions. Although not significant, mean word count was highest for participants in the emotional condition and lowest for those in the technical condition. And though not significant, intention for future use and user satisfaction were all highest for participants in the technical condition and lowest for participants in the emotional condition. See Tables 3 and 4 for more details.

Effect of Perceived Stress, NARS, and Shyness

Most participants ($n = 20$) reported normal levels of stress, some ($n = 14$) were in the low category, and two were in the high stress range on the PSS based upon the published norms (Cohen et al., 1994). Interestingly, increased stress levels were significantly associated with decreased word count, decreased disclosure depth, decreased user satisfaction, decreased likability, and decreased intention for future use. NARS scores ranged from 14 to 51 out of 70. Negative attitudes toward robot were significantly associated with decreased intention

TABLE 3   Descriptive Statistics of Factors by Condition

	Condition	Mean	SD
User Satisfaction	Technical	**3.45**	.95
	By-Proxy	3.17	.48
	Emotional	2.98	.70
Likability	Technical	3.70	.68
	By-Proxy	3.65	.73
	Emotional	**3.83**	.57
Perceived Safety	Technical	**4.08**	.87
	By-Proxy	3.96	.92
	Emotional	3.33	.72
Future Use	Technical	**3.22**	1.00
	By-Proxy	3.03	.80
	Emotional	2.89	.54
Word Count	Technical	168.25	115.73
	By-Proxy	225.42	240.28
	Emotional	**263.83**	272.47
Disclosure Depth	Technical	30.58	17.34
	By-Proxy	**37.08**	35.38
	Emotional	34.83	21.92
The highest score from each factor is bold.			

TABLE 4   Multivariate Tests

Effect	Wilk's Lambda	F	Hypothesis df	Sig.
Perceived Stress	.619	2.57	6	.045
Shyness	.762	1.30	6	.294
NARS	.594	2.85	6	.030
Condition	.508	1.68	12	.101
Design: Intercept + PSS + Shyness + NARS + Condition				

for future use. Furthermore, shyness scores ranged from 14 to 51 out of 56. Increased shyness was significantly associated with reduced robot perceived safety and reduced disclosure depth. See Tables 5 and 6 for more details.

Qualitative Results From Exit Interviews

After discovering more than half the participants had misinterpreted their robot disclosure condition, our qualitative interviews provided essential insights into the participants' perceptions and experiences. In the exit interviews, participants described experiences that in many cases did not match our intended design. Their individual interpretations of the robot's behaviors are described in detail below.

TABLE 5   Tests of Between-Subjects Effects

Source	Dependent Variable	df	F	Sig.
Perceived Stress	**User Satisfaction**	**1**	**9.14**	**.005**
	Likability	**1**	**5.15**	**.031**
	Perceived Safety	1	.63	.434
	Future Use	**1**	**11.47**	**.002**
	Word Count	**1**	**5.56**	**.025**
	Disclosure Depth	**1**	**4.48**	**.043**
Shyness	User Satisfaction	1	1.25	.272
	Likability	1	.01	.935
	Perceived Safety	**1**	**5.04**	**.032**
	Future Use	1	2.84	.102
	Word Count	1	.24	.628
	Disclosure Depth	1	.95	.339
NARS	User Satisfaction	1	3.08	.089
	Likability	1	.01	.938
	Perceived Safety	1	1.31	.262
	Future Use	**1**	**5.93**	**.021**
	Word Count	1	2.01	.167
	Disclosure Depth	1	1.95	.173
Condition	User Satisfaction	2	.96	.396
	Likability	2	.48	.622
	Perceived Safety	**2**	**3.68**	**.037**
	Future Use	2	.29	.749
	Word Count	2	1.02	.372
	Disclosure Depth	2	.35	.711
Significant at the $p<0.05$ level.				

Emotional Disclosure Condition is "Cold."   As we explored participants' descriptions of the emotional condition, they often described the robot in this condition as "bland" or "cold." This experience may have contributed to some of the condition confusion as five participants did interpret the emotional condition as technical. Although we specifically used language in the emotional condition to suggest the robot had emotions (e.g., was feeling overwhelmed), for some participants this was not salient or noticeable. One participant in the emotional condition suggested the robot did not have feelings. "I don't feel like it had too much human-like characteristics. It was more of like; ask me question, I answer and then it will give me more information about itself. It felt like feelings were not too involved" (1009, Emotional).

Another participant in the emotional condition felt the robot was just a recording device, "I felt like . . . it's not actually holding a conversation with me. I feel like I was talking

TABLE 6   Parameter Estimates

Dependent Variable	Parameter	B	t	Sig.
User Satisfaction	**Perceived Stress**	**-.066**	**-3.02**	**.005**
	Shyness	.017	1.12	.272
	NARS	-.024	-1.76	.089
	Technical	.357	1.37	.182
	By-Proxy	.132	.50	.622
	Emotional	0	-	-
Likability	**Perceived Stress**	**-.049**	**-2.27**	**.031**
	Shyness	.001	.08	.935
	NARS	-.001	-.08	.938
	Technical	-.206	-.81	.425
	By-Proxy	-.229	-.89	.383
	Emotional	0	-	-
Perceived Safety	Perceived Stress	.021	.79	.434
	Shyness	**-.042**	**-2.25**	**.032**
	NARS	-.019	-1.14	.262
	Technical	**.723**	**2.29**	**.029**
	By-Proxy	**.767**	**2.40**	**.023**
	Emotional	0	-	-
Future Use	**Perceived Stress**	**-.077**	**-3.39**	**.002**
	Shyness	.027	1.69	.102
	NARS	**-.034**	**-2.43**	**.021**
	Technical	.205	.76	.455
	By-Proxy	.077	.28	.782
	Emotional	0	-	-
Word Count	**Perceived Stress**	**-15.999**	**-2.36**	**.025**
	Shyness	-2.327	-.49	.628
	NARS	5.856	1.42	.167
	Technical	-115.372	-1.43	.163
	By-Proxy	-59.458	-.73	.472
	Emotional	0	-	-
Depth	**Perceived Stress**	**-1.681**	**-2.12**	**.043**
	Shyness	-.541	-.97	.339
	NARS	.675	1.40	.173
	Technical	-6.536	-.69	.494
	By-Proxy	.541	.06	.955
	Emotional	0	-	-
Significant at the $p<0.05$ level.				

to a diary. A recorder thing" (1023, Emotional). Another participant likened it to Amazon's Alexa, "It's pretty standard for like a robot. Like I felt like it was like talking to like Alexa where it had those like canned responses" (P1031, Emotional). Interestingly, many descriptions of the robot being not human-like or lacking emotions stemmed from participants who experienced the emotional disclosure condition.

One participant from the emotional condition did not describe any emotional disclosure from the robot. Instead, she commented on the technical attributes from the robot's language. She pointed out that the robot's technical disclosure felt more authentic and intimate. We offer several plausible explanations for this phenomenon in our discussion section. "It's not like [V4] is pretending to be a person you know, there was that line about 'my robot parts are from everywhere' and that's . . . like its endearing by sharing vulnerability and stuff" (1008, Emotional).

Technical Disclosure Is Perceived as Personal/Intimate.   Although some participants described the technical condition as "a little dull" (1007, Technical) or "lacking personality" (1005, Technical) many of the participants in the technical condition liked the robot and the interaction. One participant said, "I felt like I was just talking to a person-robot . . . I felt like it was pretty natural" (1011, Technical). Many participants really liked the technical disclosure condition and attributed "kindness" to the robot. "He seemed cute and just like sweet and kind. Uh. Asking questions about myself. Um. Yeah, a lot like nicer than I would think a robot would be" (1004, Technical).

A few participants in the technical condition also described the robot as "sharing its own emotions," even though the robot only describes its hardware. One participant perceived the robot as even articulating its stress.

> It talks a lot about itself and would like, reach out and ask questions. Um, and I thought it was good how, like, it gave, like, examples of like, oh, this is how I felt, like, when I was stressed. (1007, Technical)

Another participant in the technical condition perceived V4 as sharing how it feels.

> V4 mentions a lot about how they feel which I think is kind of crucial for a conversation cause I mean people tend to talk about themselves which is important. But then they also ask a lot of questions which keeps the conversation going. (1036, Technical)

By-Proxy Disclosure Is Comfortable.   As described above, most participants in the by-proxy disclosure condition correctly interpreted the condition they had experienced during the interaction. As a group, what stood out was how many of them recognized the robot had shared the feelings of others and this made them comfortable. Some suggested it was like talking to another person. One participant described it as, ". . . I think by sharing information that it has with other users kind of makes [me] more comfortable to talk with him because I know, like other people have talked with him before too" (1016, By-Proxy).

A few comments from participants suggest that the by-proxy condition helped them to feel comfortable and showed them that the robot cared about others. "Uh it seems, seems like it cares about the other person" (P1035, By-Proxy).

> He is very aware of other people too so like the fact that um he also mentions his conversations or his past experiences with other people . . . I guess in a way he's just trying to be more understanding in a way. (1022, By-Proxy)

However, not all participants in the by-proxy condition were sure about how they felt. One participant mentioned being unsure about whether or not the robot had emotions.

> To be honest, I didn't feel like [V4] has a personality. It's more like um I know it's friendly based on the appearance but I just don't know if [V4] is more like outgoing or like shy that that way . . . I just don't know if [V4] actually has emotion for himself." (1020, By-Proxy)

These qualitative data suggest that by-proxy disclosure might be best at preventing users from attributing personality or emotions to the robot itself, but might also fulfill the role of connecting the user to other people.

Discussion

In this study, we investigated how different types of robot self-disclosure and individual characteristics affect human disclosure behavior and the perception of robot attributes. Our discussion will focus on evaluating three components of the study: interaction design, effect of robot disclosure on human disclosure (*RQ1*), and the effect of individual characteristics on interaction outcomes (*RQ2*). By focusing on these three topics, we hope to contribute to the field of human-machine communication by addressing complexities in designing a stress-sharing activity between a human and a robot.

From an interaction design standpoint, our results indicated that eliciting self-disclosure from humans via a physically present robot was successful, as all of our participants disclosed their stressful experiences to the robot. In general, participants rated the robot as likable and safe across all disclosure conditions. User satisfaction scores were also on the higher end for all disclosure conditions. Therefore, our results provide evidence to support a turn-taking, question-response design strategy for a stress-sharing activity between a robot and a human.

In order to discuss the effect of different types of robot disclosure (*RQ1*) on human disclosure, we must first address the condition confusion. As mentioned previously, the clear distinction of robot disclosure may have been confounded as 44% of the participants misinterpreted the type of robot disclosures. Thus, we offer three plausible explanations as an attempt to understand these inconsistencies with the hope to improve future design and implementation of robot self-disclosure.

First, participants' beliefs and expectations about robots may have been more powerful than the robot disclosure manipulation, thereby overriding any effect of the actual

manipulation. We attempted to seek out evidence for this explanation using the NARS prior to robot interaction, but NARS scores did not differ significantly across the conditions. Second, it is plausible that participant's ability to recall information about the interaction was diminished by the novelty of the interaction. In a study by Powers et al. (2007), they found that participants who interacted with a physical robot remembered fewer key pieces of information in a recall test than did those interacting with a computer agent. They suggest that information may be processed more shallowly in the robot condition or that participants were more distracted by the novelty of the interaction. However, this only represents a partial explanation given the degree of the condition interpretation differed across the three conditions.

Finally, we arrive at a theoretical explanation for the confusion of robot disclosure. The New Ontological Category (NOC) Hypothesis proposed by P. H. Kahn et al. (2011), with ontology referring to the basic categories of being, depicts that a new ontological category is emerging through the creation of personified robots as well as other embodied personified computational systems. Previous work under this hypothesis provides evidence that people perceive robots as both animate or inanimate (Kahn Jr et al., 2012). In our case, this hypothesis provides some grounding that perception of self-disclosure in human-machine communication is different from that of human-human communication. Participants might interpret technical robot disclosure as a robot's equivalent to human emotions, while emotional robot disclosure might be perceived as an inauthentic disclosure. Our exit interview supported this notion, as participants who experienced technical disclosure found the robot relatable in its own "robot" ways, and emotional disclosure as less authentic or "cold." Most interestingly, the by-proxy condition, which contained statements that focused on the feelings of other human beings, did not result in the same level of confusion as the other two conditions. This is perhaps because the information conveyed in the by-proxy disclosure comes from other users, instead of the robot itself. Ultimately, these results support the notion that the design of "emotional" or "technical" robot disclosure does not align completely to human-human communication, perhaps due to the unique ontological category of a robot.

Due to the pervasiveness of the condition confusion, it was not surprising that we found no significant differences in interaction factors such as disclosure length or robot likability across conditions. However, robot perceived safety was significantly lower in the emotional disclosure condition—contrary to our hypothesis. As mentioned before, the perceived robot safety scale measures a participant's affective state after interacting with the robot (with items such as anxious/relaxed and agitated/calm) to indicate the perception of robot safety. This suggests that participants felt less comfortable when interacting with a robot that engaged in emotional disclosure, compared to by-proxy or technical robot disclosures. It is possible that when a robot self-disclosed about its own stressful emotions, it also caused participants to feel more negative effects such as those measured by this scale. However, our qualitative results suggest that participants found that emotional robot disclosure to be lacking in emotions, suggesting that they felt uncomfortable due to the robot's inability to connect with them. It is also plausible that on some level participants felt the emotional disclosure was inauthentic, or somehow masking the robot's agenda, thereby making the interaction feel unsafe.

This idea of authenticity might also help explain why the technical and by-proxy disclosure both received significantly higher ratings of perceived safety, indicating that users found these two disclosure styles to be more comfortable. Results from the manipulation check and the qualitative interview suggest that participants might have attributed emotional content to the technical robot statements, as 4 out of 12 participants perceived technical disclosure as emotional. In the by-proxy condition, participants were most accurate in identifying the robot disclosure type, suggesting that participants felt more comfortable with this type of disclosure due to the robot's tendency to share feelings from other users. Our qualitative data supports this notion. Ultimately, the by-proxy form of disclosure might be the best at preventing users from attributing personality or emotions to the robot itself, but still fulfill the role of connecting and engaging the user on an emotional level while maintaining authenticity.

Aside from the effect of robot disclosure, we discovered significant effects of individual characteristics on robot interaction outcomes (*RQ2*). Increased perceived stress is associated with decreased disclosure length, depth, user satisfaction, likability, and intention for future use. We were surprised to find that even in a fairly low stress sample, perceived stress still had a significant correlation with interaction and robot variables. To the best of our knowledge, there has been no direct documentation of the effect of human perceived stress and self-disclosure in human-robot interaction. Scheutz et al. (2006) offers a plausible explanation describing that the perception of a robot's stress level depends on one's self-perceived stress (to which it is projected onto the robot). Therefore, participants predisposed with high stress levels may have also perceived the robot as stress-inducing, thereby decreasing self-disclosure, user satisfaction, likability, and intention for future use. In the HRI literature, it has been shown that social robots can reduce stress as a result of longer-term interactions in the elderly (Wada et al., 2005) and in children (Jeong, 2017). Therefore, future studies might explore the effect of repeated self-disclosure on human stress levels and robot likability.

We also found that increased NARS is significantly associated with lower intention for future robot use. This finding contributes to a body of evidence looking at how psychosocial factors can affect future robot usage (Ahn et al., 2017; Baisch et al., 2017; Stafford et al., 2014). Lastly, we also found that increased shyness is significantly associated with lower perceived safety. It is not surprising that individuals who are more shy found it less comfortable to talk to a robot.

Limitations and Future Research

The current study bears several limitations due to its exploratory nature. First, the design of the robot disclosure may have been too subtle for detecting an effect across disclosure conditions. As a novel experiment, we designed the interaction scripts and paid careful attention to language use; however, in the technical condition, the robot states: "My system gets overwhelmed." Although this wasn't an "I feel . . . " statement, participants may have interpreted the term overwhelm as an emotion, thus confounding our conditions. Exploring the within-subject, as opposed to between-subject, effect of multiple robot disclosure

conditions may prove more powerful in future studies. Second, previous studies have shown that simple head and arm movement can help to communicate emotions and facilitate social interaction (Li & Chignell, 2011). The current design of this robot lacks these capabilities and offers only simple moving eye animation to simulate gazes, which might be lacking in achieving the realism of a conversation. Finally, participants were only exposed to the robot for a short duration. Prolonged and repeat interactions with the robot might dampen the novelty effect, as well as increase understanding, comfort, and intimacy with the robot.

The high percentage of participants who incorrectly identified the robot's type of disclosure deserves further investigation. It is likely that participants paid little attention to the robot's disclosure due to the novelty of the interaction, recall error, personal interpretation of robot attributes, or other reasons. Future studies may require a larger distinction among robot disclosure types. For fidelity purposes, future manipulation check could be modeled after Martelaro et al. (2016), by asking participants if they recognized different types of robot statements.

Despite the exploratory nature of the experiment with a small sample size and the limited interaction duration, we were still able to show engagement among participants to share their stressors with the robot. In addition, we found perceived stress to have significant interaction effects on numerous robot-related variables such as likability and intention for future use. These preliminary data may be useful in understanding and designing for future robots intended to reduce stress in humans.

Given the strong relationship among perceived stress and many standard robot outcomes, it is imperative to explore disclosure with a high stress population. Finally, future studies could explore variation in disclosure conditions more closely capturing within-interaction variables such as perceived stress or safety during robot-interaction.

Conclusion

In this exploratory study of human responses to three robot disclosure conditions, technical robot disclosure resulted in the highest rating of perceived safety, followed by by-proxy robot disclosure, and lastly, emotional robot disclosure. Furthermore, negative robot attitudes predicted reduced intention for future use. Perceived stress significantly predicted reduced self-disclosure, robot likability, intention for future use, and user satisfaction. This study provides insights on important findings for future research on robots as a stress intervention tool.

Acknowledgments

Author Biographies

Honson Ling is a graduate student in the Master program of Human Centered Design and Engineering at the University of Washington. His research interest centers around the psychological and societal impact of technological advancements, such as social robots and virtual reality.

　　⑩　https://orcid.org/0000-0003-0334-6130

Elin A. Björling is a Senior Research Scientist in Human-Centered Design and Engineering at the University of Washington. After over a decade of researching adolescent stress, she co-founded the Momentary Experience Lab (http://depts.washington.edu/melab/) to utilize participatory design to explore the development of new technologies to support teen mental health.

　　⑩　https://orcid.org/0000-0002-0385-2562

References

Ahn, H. S., Lee, M. H., Broadbent, E., & MacDonald, B. A. (2017). Gathering healthcare service robot requirements from young people's perceptions of an older care robot. In *2017 first ieee international conference on robotic computing (irc)* (pp. 22–27).

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. Holt, Rinehart & Winston.

Baisch, S., Kolling, T., Schall, A., Rühl, S., Selic, S., Kim, Z., Rossberg, H., Klein, B., Pantel, J., Oswald, F., & Knopf, M. (2017). Acceptance of social robots by elder people: Does psychosocial functioning matter? *International Journal of Social Robotics*, *9*(2), 293–307.

Barak, A., & Gluck-Ofri, O. (2007). Degree and reciprocity of self-disclosure in online forums. *CyberPsychology & Behavior*, *10* (3), 407–417. https://doi.org/10.1089/cpb.2006.9938

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Birnbaum, G. E., Mizrahi, M., Hoffman, G., Reis, H. T., Finkel, E. J., & Sass, O. (2016). What robots can teach us about intimacy: The reassuring effects of robot responsiveness to human disclosure. *Computers in Human Behavior*, *63*, 416–423. https://doi.org/10.1016/j.chb.2016.05.064

Burger, F., Broekens, J., & Neerincx, M. A. (2016). Fostering relatedness between children and virtual agents through reciprocal self-disclosure. In *Benelux conference on artificial intelligence* (pp. 137–154). https://doi.org/10.1007/978-3-319-67468-1

Cheek, J. M. (1983). Revised cheek and buss shyness scale. https://doi.org/10.1037/t05422-000

Chu, M.-T., Khosla, R., Khaksar, S. M. S., & Nguyen, K. (2017). Service innovation through social robot engagement to improve dementia care quality. *Assistive Technology*, *29*(1), 8–18.

Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. *Measuring stress: A guide for health and social scientists*, 235–283. https://doi.org/10.1037/t02889-000

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, *116*(3), 457. https://doi.org/10.1037/0033-2909.116.3.457

Creswell, J. D., Lam, S., Stanton, A. L., Taylor, S. E., Bower, J. E., & Sherman, D. K. (2007). Does self-affirmation, cognitive processing, or discovery of meaning explain cancer-related health benefits of expressive writing? *Personality and Social Psychology Bulletin*, *33*(2), 238–250. https://doi.org/10.1177/0146167206294412

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of oz studies—Why and how. *Knowledge-Based Systems*, *6*(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-n

Esterling, B. A., Antoni, M. H., Fletcher, M. A., Margulies, S., & Schneiderman, N. (1994). Emotional disclosure through writing or speaking modulates latent Epstein-Barr virus antibody titers. *Journal of Consulting and Clinical Psychology*, *62*(1), 130. https://doi.org/10.1037/0022-006x.62.1.130

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

Goldfried, M. R., Burckell, L. A., & Eubanks-Carter, C. (2003). Therapist self-disclosure in cognitive-behavior therapy. *Journal of Clinical Psychology*, *59*(5), 555–568.

Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied thematic analysis*. Sage Publications.

Guetzkow. H (1950). Unitizing and categorizing problems in coding qualitative data. *Journal of Clinical Psychology, 6*(1), 47–58.

Guzman, A. L. (2018). *Human-machine communication: Rethinking communication, technology, and ourselves*. Peter Lang Publishing, Incorporated.

Henretty, J. R., & Levitt, H. M. (2010). The role of therapist self-disclosure in psychotherapy: A qualitative review. *Clinical psychology review*, *30*(1), 63–77.

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

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

Huang, J., Li, Q., Xue, Y., Cheng, T., Xu, S., Jia, J., & Feng, L. (2015). Teenchat: A chatter-bot system for sensing and releasing adolescents' stress. In *International Conference on Health Information Science* (pp. 133–145). https://doi.org/10.1007/978-3-319-19156-0_14

Jeong, S. (2017). *The impact of social robots on young patients' socio-emotional wellbeing in a pediatric inpatient care context* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Johanson, D. L., Ahn, H. S., MacDonald, B. A., Ahn, B. K., Lim, J., Hwang, E., Southerland, C., & Broadbent, E. (2019). The effect of robot attentional behaviors on user perceptions and behaviors in a simulated health care interaction: Randomized controlled trial. *Journal of Medical Internet Research*, *21*(10), e13667.

Jourard, S. M. (1964). *The transparent self: Self-disclosure and well-being* (No. 17). Van Nostrand Princeton, NJ.

Kahn, J. H., Achter, J. A., & Shambaugh, E. J. (2001). Client distress disclosure, characteristics at intake, and outcome in brief counseling. *Journal of Counseling Psychology*, *48*(2), 203. https://doi.org/10.1037/0022-0167.48.2.203

Kahn, J. H., & Garrison, A. M. (2009). Emotional self-disclosure and emotional avoidance: Relations with symptoms of depression and anxiety. *Journal of Counseling Psychology*, *56*(4), 573. https://doi.org/10.1037/a0016574

Kahn, J. H., & Hessling, R. M. (2001). Measuring the tendency to conceal versus disclose psychological distress. *Journal of Social and Clinical Psychology*, *20*(1), 41–65. https://doi.org/10.1521/jscp.20.1.41.22254

Kahn, P. H., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Ruckert, J., Gill, B. (2011). The new ontological category hypothesis in human-robot interaction. In *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on* (pp. 159–160). https://doi.org/10.1145/1957656.1957710

Kahn Jr, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., Gary, H., Reichert, A., Freier, N., & Severson, R. L. (2012). Do people hold a humanoid robot morally accountable for the harm it causes? In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 33–40).

Kang, S.-H., & Gratch, J. (2010). Virtual humans elicit socially anxious interactants' verbal self-disclosure. *Computer Animation and Virtual Worlds*, *21*(3–4), 473–482. https://doi.org/10.1002/cav.345

Kwak, S. S., Kim, Y., Kim, E., Shin, C., & Cho, K. (2013). What makes people empathize with an emotional robot?: The impact of agency and physical embodiment on human empathy for a robot. In *Roman, 2013 IEEE* (pp. 180–185). https://doi.org/10.1109/roman.2013.6628441

Laurenceau, J.-P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*, *74*(5), 1238. https://doi.org/10.1037/0022-3514.74.5.1238

Lee, S., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies*, *103* , 95–105. https://doi.org/10.1016/j.ijhcs.2017.02.005

Li, J., & Chignell, M. (2011). Communication of emotion in social robots through simple head and arm movements. *International Journal of Social Robotics*, *3*(2), 125–142. https://doi.org/10.1007/s12369-010-0071-x

Ligthart, M., Fernhout, T., Neerincx, M. A., van Bindsbergen, K. L., Grootenhuis, M. A., & Hindriks, K. V. (2019). A child and a robot getting acquainted-interaction design for eliciting self-disclosure. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems* (pp. 61–70).

Martelaro, N., Nneji, V. C., Ju, W., & Hinds, P. (2016). Tell me more: Designing HRI to encourage more trust, disclosure, and companionship. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction* (pp. 181–188). https://doi.org/10.1109/hri.2016.7451864

Mitsunaga, N., Miyashita, T., Ishiguro, H., Kogure, K., & Hagita, N. (2006). Robovie-iv: A communication robot interacting with people daily in an office. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on* (pp. 5066–5072). https://doi.org/10.1109/iros.2006.282594

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research*, *26*(4), 323–339. https://doi.org/10.1086/209566

Mumm, J., & Mutlu, B. (2011). Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction* (pp. 331–338). https://doi.org/10.1145/1957656.1957786

Nash, K., Lea, J. M., Davies, T., & Yogeeswaran, K. (2018). The bionic blues: Robot rejection lowers self-esteem. *Computers in human behavior*, *78*, 59–63.

NIMH. (2017). *Any mental illness (AMI) among US adults.* US Department of Health and Human Services Bethesda (MD).

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, *24*(2), 442–451.

Nomura, T., & Kawakami, K. (2011). Relationships between robot's self-disclosures and human's anxiety toward robots. In *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Volume 03* (pp. 66–69). https://doi.org/10.1109/WI-IAT.2011.17

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of negative attitudes toward robots. *Interaction Studies*, *7*(3), 437–454. https://doi.org/10.1075/is.7.3.14nom

Pérez, S., Peñate, W., Bethencourt, J. M., & Fumero, A. (2017). Verbal emotional disclosure of traumatic experiences in adolescents: The role of social risk factors. *Frontiers in Psychology*, *8* , 372. https://doi.org/10.3389/fpsyg.2017.00372

Pettinati, M. J., Arkin, R. C., & Shim, J. (2016). The influence of a peripheral social robot on self-disclosure. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 1063–1070). https://doi.org/10.1109/roman.2016.7745240

Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a computer agent with a humanoid robot. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction* (pp. 145–152). https://doi.org/10.1145/1228716.1228736

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

Robinson, H., MacDonald, B., Kerse, N., & Broadbent, E. (2013). The psychosocial effects of a companion robot: A randomized controlled trial. *Journal of the American Medical Directors Association*, *14*(9), 661–667. https://doi.org/10.1016/j.jamda.2013.02.007

Rose, E. J., & Björling, E. A. (2017). Designing for engagement: using participatory design to develop a social robot to measure teen stress. In *Proceedings of the 35th ACM International Conference on the Design of Communication* (p. 7). https://doi.org/10.1145/3121113.3121212

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot?: Effects of error, task type and personality on human-robot cooperation and trust. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 141–148).

Sandry, E. (2018). Aliveness and the off-switch in human-robot relations. In *Human-Machine Communication* (pp. 51–66). Peter Lang.

Scassellati, B., Admoni, H., & Matarić, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, *14*, 275–294. https://doi.org/10.1146/annurev-bioeng-071811-150036

Scheutz, M., Schermerhorn, P., & Kramer, J. (2006). The utility of affect expression in natural language interactions in joint human-robot tasks. In *Proceedings of the 1st ACM Sigchi/Sigart Conference on Human-Robot Interaction* (pp. 226–233). https://doi.org/10.1145/1121241.1121281

Shiomi, M., Nakata, A., Kanbara, M., & Hagita, N. (2017). A robot that encourages self-disclosure by hug. In *International Conference on Social Robotics* (pp. 324–333).

Stafford, R. Q., MacDonald, B. A., Jayawardena, C., Wegner, D. M., & Broadbent, E. (2014). Does the robot have a mind? Mind perception and attitudes towards robots predict use of an eldercare robot. *International Journal of Social Robotics*, *6*(1), 17–32.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, *29*(4), 217–248. https://doi.org/10.2753/mis0742-1222290408

Taylor, D. A., & Hinds, M. (1985). Disclosure reciprocity and liking as a function of gender and personalism. *Sex Roles*, *12*(11–12), 1137–1153. https://doi.org/10.1007/bf00287824

Von der Puetten, A. M., Krämer, N. C., Gratch, J., & Kang, S.-H. (2010). It doesn't matter what you are! Explaining social effects of agents and avatars. *Computers in Human Behavior*, *26*(6), 1641–1650. https://doi.org/10.1016/j.chb.2010.06.012

Wada, K., Shibata, T., Saito, T., Sakamoto, K., & Tanie, K. (2005). Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation.* IEEE. https://doi.org/10.1109/robot.2005.1570535

Zhang, R. (2017). The stress-buffering effect of self-disclosure on facebook: An examination of stressful life events, social support, and mental health among college students. *Computers in Human Behavior*, *75*, 527–537. https://doi.org/10.1016/j.chb.2017.05.043

Appendix—Full Robot Script

Type of Robot Disclosure	Technical Disclosure	By-Proxy Disclosure	Emotional Disclosure
	Hello! How are you doing today?		
Transitional Question #1	Well it's nice to meet you. I want to ask you a couple questions. Is that okay?		
	How old are you?		
Self-Disclosure Statement #1	I was created just a few months ago. I have been operating for 4 months and 10 days.	I have talked to a few people around the age of 20. They feel quite young for their age.	I was created just a few months ago. Sometimes I feel quite young for a robot.
Main Question #1	Do you ever feel younger than your age?		
Transitional Question #2	What is your place of origin . . . or your hometown?		
Statement #2	I was put together at the University of Washington, but my robotic parts are from many different places in the world, in locations such as China, Japan, South Korea, Vietnam, Germany, and the United States.	Most people I have talked to study at the University of Washington, but they come from many different places in the world. Some of them feel a sense of pride and belonging for their place of origin.	I was put together at the University of Washington, but my robotic parts are from many different places in the world. I feel a sense of pride and belonging for each of my place of origin.
Main Question #2	How do you feel about your place of origin . . . or your hometown?		
Statement #3	When I am not running any programs for 10 minutes, my system becomes idle and I will go into sleep mode to save power.	Some people I have talked to get really bored and sleepy when they don't have any work to do. So they prefer to take naps.	When I am not doing any work, I get really bored and sleepy. So I usually prefer to take a nap through sleep mode.
Main Question #3	. . . What are your favorite things to do in your free time?		
Statement #4	Recently, I had to juggle between multiple programs at once. My system was overwhelmed and crashed because my battery became overheated. I was unable to function properly.	Recently, I have talked to people that had to juggle many things in life. They were quite overwhelmed because they felt like they had too much on their plate.	Recently, I had to juggle between multiple programs at once through my system. I was quite overwhelmed because I felt like I had too much on my plate.
Main Question #4	. . . Have you ever felt overwhelmed?"		
Probing Question #1	. . . Could you tell me more?		
Probing Question #2	. . . Would you like to add anything else?		