PRIFYSGOL BANGOR / BANGOR UNIVERSITY

Building Long-Term Human–Robot Relationships: Examining Disclosure, Perception and Well-Being Across Time
Laban, Guy; Kappas, Arvid; Morrison, Val; Cross, Emily

International Journal of Social Robotics

Publisher's PDF, also known as Version of record

Cyswllt i'r cyhoeddiad / Link to publication

28. Apr. 2024

Building Long-Term Human–Robot Relationships: Examining Disclosure, Perception and Well-Being Across Time

Guy Laban[1,2] · Arvid Kappas[3] · Val Morrison[4] · Emily S. Cross[2,5,6]

Abstract

While interactions with social robots are novel and exciting for many people, one concern is the extent to which people's behavioural and emotional engagement might be sustained across time, since during initial interactions with a robot, its novelty is especially salient. This challenge is particularly noteworthy when considering interactions designed to support people's well-being, with limited evidence (or empirical exploration) of social robots' capacity to support people's emotional health over time. Accordingly, our aim here was to examine how long-term repeated interactions with a social robot affect people's self-disclosure behaviour toward the robot, their perceptions of the robot, and how such sustained interactions influence factors related to well-being. We conducted a mediated long-term online experiment with participants conversing with the social robot Pepper 10 times over 5 weeks. We found that people self-disclose increasingly more to a social robot over time, and report the robot to be more social and competent over time. Participants' moods also improved after talking to the robot, and across sessions, they found the robot's responses increasingly comforting as well as reported feeling less lonely. Finally, our results emphasize that when the discussion frame was supposedly more emotional (in this case, framing questions in the context of the COVID-19 pandemic), participants reported feeling lonelier and more stressed. These results set the stage for situating social robots as conversational partners and provide crucial evidence for their potential inclusion in interventions supporting people's emotional health through encouraging self-disclosure.

Keywords Social robots · Human-robot interaction · Longitudinal · Well-being · Self-Disclosure · Social perception

✉ Emily S. Cross
ecross@ethz.ch

Guy Laban
gl538@cam.ac.uk; guy.laban@glasgow.ac.uk

Arvid Kappas
akappas@constructor.university

Val Morrison
v.morrison@bangor.ac.uk

[1] Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

[2] School of Psychology and Neuroscience, University of Glasgow, Glasgow, UK

[3] School of Business, Social and Decision Sciences, Constructor University, Bremen, Germany

[4] Department of Psychology, Bangor University, Bangor, UK

[5] MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, NSW, Australia

[6] Professorship for Social Brain Sciences, ETH Zürich, Zürich, Switzerland

1 Introduction

Social robots have been shown to effectively elicit socially meaningful behaviours and emotions from humans across a number of experimental and real-world contexts [1–3]. Nevertheless, one of the challenges to human–robot interaction (HRI) research is employing and evaluating long-term interactions, especially in people's natural settings. Since interactions with social robots are novel and exciting for many people, one particular concern in this specific area of HRI is the extent to which behavioural and emotional expressions might develop from initial interactions with a robot, when its novelty is particularly salient, to responses, behaviours, and perceptions that are sustained over time [4, 5]. Empirical studies in this area of HRI research are often limited to controlled laboratory settings, due to various logistical (e.g., limited number of robots per lab and robots' high cost) and technical factors (e.g., multiple computers or other controlling devices required to coordinate a robot's behaviours and/or requirements for skilled Wizard-of-Oz (WoZ) operation). These challenges can make it difficult

for HRI researchers to gain insights into factors that shape people's long-term interactions with social robots in natural, real-world settings. This is especially noticeable with studies that are focused on evaluating the use and utility of social robots in social settings. Robots for these settings are often designed to interact and communicate with humans or other agents (such as pets or other robots) by following social scripts and rules relevant to their role and function within a given social setting [1, 6]. Our understanding of social robots' potential scope and limitations will be substantially informed via experiments where people can interact across longer periods of time with robots in natural social settings, such as within one's home, workplace, local clinic, or school.

Overcoming this challenge will be particularly important when devising interactions to support people's well-being. Social robots are widely studied and are gradually being introduced in care settings, aimed at supporting people's physical and mental well-being [1]. However, due to the complexity of administering social robotic interventions, studies in the field rarely establish ecologically valid interactions with human users, and instead often implement insufficient methods (e.g., using single-subject studies, quasi-experimental designs, cross-sectional research designs, etc.), or explore single interactions rather than ongoing longitudinal interventions [see 7].

Considering social robots' social features [8, 9], animated design [2, 10], and autonomous abilities [1], social robots situated in people's homes hold potential for helping to monitor physical health, as well as improve emotional well-being by engaging in conversation that fosters self-disclosure. Self-disclosure is a communication behaviour aimed at introducing and revealing oneself to others, and it plays a key role in building relationships between two individuals [12, 13]. It serves an evolutionary function of strengthening interpersonal relationships, while also fostering a wide variety of health benefits, including helping people to cope with stress and traumatic events through eliciting help and support [14–16]. Moreover, self-disclosure appears to play a critical role in successful treatment outcomes [17] and has a positive impact on mental and physical health [18]. For health interventions to succeed, they depend on open channels of communication where individuals can disclose needs and emotions, from which a listener can identify stressors and respond accordingly [19, 20]. This is crucial for interventions with social robots, as human behaviour and emotions are analysed and synthesized by machines from human output, to respond and react appropriately [21].

Given the necessity of studying social robotic interventions over longer time courses than what are often seen in one-off laboratory studies, as well as the importance of self-disclosure for psychological health and HRI,

here we aimed to evaluate people's self-disclosure during interactions with a social robot over time. More specifically, we explored how prolonged and intensive interactions with a social robot affect people's self-disclosure behaviour toward the robot, perceptions of the robot, and additional factors related to well-being. Therefore, we were asking—

RQ1: To what extent are people's self-disclosures, perceptions of the robot, and well-being affected over time during long-term interactions with a social robot?

To build a more complete understanding of the application of social robots in different emotional settings, we were also interested in the role of the interaction's discussion frame. Hence, we were also asking—

RQ2: To what extent are people's self-disclosures, perceptions of the robot, and well-being, affected due to the discussion frame during long-term interactions with a social robot?

To address these research questions, we conducted a mediated long-term online experiment with participants conversing with a social robot 10 times over 5 weeks about general everyday topics. Participants were allocated to two groups, discussing topics framed in the context of the Covid-19 pandemic, or the same topic except the discussion had no explicit mention of the Covid-19 pandemic.

2 Related Work

The mere-exposure effect refers to the psychological phenomenon where people develop a preference for things they are repeatedly exposed to [22]. In the context of long-term HRI, the mere-exposure effect operates differently compared to the realm of Human–Computer Interaction (HCI) focusing on usability. In HCIs, users often acquire a positive attitude towards tools and objects through repeated use and familiarity, leading to improved usability [23]. However, in the domain of HRI, where social interaction with robots is a key component [1], the dynamics change. The social dynamic in HRI sets it apart from traditional HCI. Unlike HCI, where the focus is primarily on optimizing usability and functionality [23, 24], HRI incorporates a social component, aiming to create a sense of companionship or collaboration [e.g., 25]. Robots are designed to engage with humans in a more social manner, simulating human-like behaviours, gestures, and communication. This social element introduces a unique dynamic, where humans naturally seek to establish social connections, rapport, and even attribute human-like qualities to the robots [1, 9]. The social dynamics in HRI require a deeper understanding of human–robot communication beyond usability, and instead focus on the development of a social bond between humans and robots.

Longitudinal designs are important for understanding people's long-term adaptation of social robots, and moreover, to further understand human behaviour and perception of social robots and how it changes over time [26]. While single interaction studies provide us with interesting insights into human behaviour when engaging with robots, we are often challenged to learn from these studies as "in the wild" application of robots aims to develop machines that people interact with over sustained periods of time [1, 27]. One of the most significant (and common) limitations to one-off HRI studies relates to novelty effects [see 4, 5], while long-term studies have often found evidence for reduced engagement with various robotic platforms over time [26, 28]. As social robots are a new emerging technology that is still novel and exciting for most, users often have high expectations for social robots and experience dissonance when a social robot's performance fails to meet their expectations. Accordingly, when users interact with robots over time, they tend to perceive them as being less social as interactions go on as their expectations of the robot are not being fulfilled [5]. Previous studies show that even household robotic devices that are not particularly social (like the Roomba vacuum cleaner) suffer from the novelty effect [29], with users being excited about the robotic device at first and using it less as they get familiar with it.

Due to the constrained and highly choreographed nature of many HRI studies, deep insights into people's responses and interactions with robots in natural settings remain relatively rare. Of the field studies that have conducted HRI research in these spaces, important insights are emerging from both single interaction [e.g., 30, 31] and repeated interaction [e.g., 32–38] studies, with much of this work taking place in public spaces or tied to specific settings like education [e.g., 36, 39–43], care [e.g., 37, 38, 44–47], or rehabilitation [e.g., 32, 35, 48–51]. Longitudinal studies that address similar questions with disembodied agents such as virtual assistants and chatbots [e.g., 52–54] benefit from access to users' personal devices, whereas research with physically embodied artificial agents (i.e., social robots) remains far rarer due to challenges with logistical and cost barriers to situating these devices in users' domestic settings (i.e., in their home environment) to explore single or repeated interactions. While several attempts have been made before to reduce the barriers to feasibility for such work [c.f., 33, 34, 38, 55, 56], we still know relatively little about user perceptions of and behaviours toward social robots when these take place in familiar home environments. Further insights into the challenges and opportunities afforded by placing social robots into familiar domestic settings should aid human-robot communication in general, as well as further refine the development and utility of these machines for commercial use.

2.1 Social Robots for Well Being

Social robots hold great potential for delivering or improving psycho-social interventions [7], supporting mental health [57], monitoring symptoms of chronic psychopathologies [58], aiding rehabilitation [35] and providing much-needed physical and social support across a number of daily life settings [1]. For example, a previous study by Nomura and colleagues [59] showed the benefits of employing social robots for minimising social tensions and anxieties, describing that participants with higher social anxiety tended to feel less anxious and demonstrate lower tensions when knowing that they would interact with robots in opposition to humans. In fact, a recent paper [60] stresses the benefits of employing social robots as interventions for social anxiety, stating that these could complement the support provided by clinicians. The authors explain that social robots could support people to get into therapy and maximize the effectiveness of the therapy by increasing the patients' engagement and continuing the support outside the therapy session. In a previous paper [58] we addressed similar benefits of using social robots for diagnosing and treating people suffering from post-traumatic stress disorder (PTSD), social robots can assist with overcoming several logistical and social barriers that trauma survivors face when required to monitor symptoms and when seeking mental health interventions.

Beyond supporting people with clinically diagnosed psychopathologies like PTSD and anxiety, social robots could also provide emotional support via self-managed interventions to healthy individuals that might experience difficult emotional situations and stressors in their daily lives. Previous studies administered the application of social robots in emotionally supportive settings showing meaningful outcomes in terms of cognitive change and affect. A study by Bodala and colleagues [47] employed a social robot delivering teleoperated mindfulness coaching for five weeks. Another example by Axelsson and colleagues [61] tested a robotic coach conducting positive psychology exercises, showing positive mood change after participation in the robotic intervention. Robotic interventions for people's well-being are rarely taking place in people's homes. One successful example is a study employing the social robot Jibo as a positive psychology coach to improve students' psychological well-being in students' on-campus housing. The study results describe a positive effect on students' psychological well-being with positive mood change, and also students expressing their motivation to change their psychological well being [38]. Other studies show positive outcomes in terms of behavioural change. A series of studies by Robinson and colleagues used social robots to deliver behaviour change interventions, applying verbal motivational interventions for reducing high-calorie snack

consumption. The studies showed promising results addressing the behavioural change using objective measurements like weight loss [see 45], and also via qualitative data addressing the subjective experiences of the participants during such interventions [see 62]. A similar intervention have been tested with a clinical population showing potential for using social robots for diabetes management [44].

The recent COVID-19 pandemic further illuminated the potential of social robots as an assistive technology in times when strict infection control measures mandate physical distancing between people. Several researchers have argued that physically embodied social robots should be able to assist with a number of tasks to help keep people physically and mentally healthy, ranging from temperature taking and food and supply delivery to providing companionship for individuals suffering from loneliness [1, 63–66], and even mediating social interactions with other individuals [67]. Nevertheless, as discussions concerning the potential applications for social robots became more prominent during the pandemic, HRI research was limited due to social distancing and the inability to use lab facilities for research that is highly dependent on laboratory-constrained environments. The pandemic forced most individuals (including researchers) to adopt computer-mediated means of communication (CMC) [68]. Following the wholesale adaptation to CMC during the pandemic, the current research sets forth a means for conducting rigorous and reproducible social robotics research to explore people's engagement with social robot-mediated interactions within their own homes. More generally, this research sets the stage for further research exploring online mediated speech-based psychosocial interventions with social robots when public health, cost, or logistical barriers prevent situating a physically embodied robot in users' homes across the long term.

2.2 Self-Disclosing to Robots and Artificial Agents

Several studies address self-disclosure to social robots in single sessions [e.g., 3, 69–73], however, few studies to date have addressed self-disclosures to robots in long-term settings [e.g., 74]. Previous studies describe that in single interactions, people's subjective perceptions of their self-disclosures to robots tend to align objectively well with their actual disclosures. Moreover, people tend to share more information with humans than with humanoid social robots or other artificial agents [3]. Yet, a different study by Nomura and colleagues [59] found that speech interactions with a social robot elicited lower tensions compared to interactions with a human agent. Another recent study explains that people might self-disclose more to a robot when conversing with a robot that changes their listening attitude [72]. Long-term studies with disem-

bodied conversational agents give us some evidence of the nature of long-term self-disclosure to artificial agents. For example, a longitudinal study by Croes and Antheunis [52] tested long-term interactions with the chatbot Mitsuku via 7 interactions that were conducted over 3 weeks. Their results show that social processes decreased after each interaction with Mitsuku and that participants reported lower feelings of friendship with Mitsuku across sessions. They describe the presence of the novelty effect, with participants describing how Mitsuku became predictable after the first session. An additional study by Croes and Antheunis [53] showed that in self-disclosure interactions despite feeling more anonymous when interacting with chatbots, people trust humans more than they trust chatbots and reported for higher degrees of social presence.

2.3 Using Self-Disclosure for a Social Robotic Intervention to Support Mental Health

The literature describes that various forms of human–human self-disclosure can support and improve mood and provide a convenient space for concealment and regulating emotions with many health benefits. For example, James Pennebaker writing disclosure paradigm [75, 76] helps people to facilitate their emotions when writing about their own experiences. Previous studies have reported that people in a bad mood benefited more from disclosing to a robot than participating in writing disclosures in a journal [70] or on social media [77]. Another good example is affect labelling, a simple and implicit emotional regulation technique aimed at explicitly expressing emotions, or in other words—putting feelings into words [78]. In addition, the act of self-disclosure is highly useful for emotional introspective process, self-reflection on one's emotions, actions, and behaviours [79], and is a meaningful act of mindfulness [80]. In the previous section, we mentioned several studies applying social robots in emotionally supportive settings [see 38, 47, 61], but social robotic interventions rarely encourage open self-disclosure [see 81]. Considering the vast evidence for the positive effect of self-disclosure on emotional well-being, our behavioural paradigm was aimed at encouraging participants to self-disclose to a social robot as a therapeutic activity. Engaging a robot in a reciprocal conversational interaction is a complex technical task, that might negatively affect people's disclosures and perceptions of the robot and the interaction due to the robot's communication limitations. However, here we suggest that employing a social robot for encouraging and listening to people's disclosure (an act that would not be as limiting to the robot's communication skills) would have a positive effect on people's disclosures and perceptions. Furthermore, we suspect that by engaging people in self-disclosures to a social robot, participants would be

engaged in affect labelling [78] and it will positively affect their well-being.

3 Methods

Consistent with recent proposals [82, 83], we pre-registered the study and report for how we determined our sample size, all data exclusions, all manipulations and all measures in the study [see 84]. In addition, following open science initiatives [e.g., 85], the de-identified data set, stimuli and analysis code associated with this study are freely available online [86]. By making the data available, we enable and encourage others to pursue tests of alternative hypotheses, as well as more exploratory analyses.

3.1 Experimental Design

A between-groups 2 (Discussion Frame: Covid-19 related or general) by 10 (chat sessions across time) repeated measures experimental design was followed. Participants were randomly assigned to one of the two discussion frames groups, according to which they conversed with the robot Pepper (SoftBank Robotics) via Zoom video chats about general everyday topics (e.g., social relationships, work-life balance, health and well-being; see Table 1) for 10 sessions. One group's conversation topics were framed within the context of the Covid-19 pandemic (e.g., social relationships during the pandemic, sustaining mental health during the pandemic, etc.), whereas the other group's conversation topics were similar, except no explicit mention of the Covid-19 pandemic was ever made (see Sect. 3.4.3). Each interaction consisted of the robot asking the participant 3 questions (x3 repetitions). The topic of each interaction was assigned randomly before the experimental procedure started, as was the order of the questions. Participants were scheduled to interact with the robot twice a week during prearranged times for five weeks.

3.2 Participants

3.2.1 Sample

A priori power calculations using G*power software [87, 88] suggest that for reasonable power (0.83) to detect small to medium effect sizes, a sample size of 22 participants would be required. Due to the relatively complex data collection procedure and the potential for a high dropout rate, we recruited 40 participants via the Prolific website. One participant dropped out, resulting in a final sample size of 39 participants. Participants were between the ages of 18 and 60 ($M = 36.41$, $SD = 12.20$), 54% identify as females, and the rest identify as males. More than half (59%) of the sample reports having a Bachelor's degree as their highest level of education, and

more than half of the sample (51.3%) are employed full-time. 55% of the sample are either married (33.3%) or in a relationship (21.7%). 41% of the sample have at least one child. Most of the participants (97.4%) did not live on their own during their participation in the study, with an average number of 3.36 individuals ($SD = 1.37$) in a household (including the participant). Almost all of the participants (92%) did not have previous experience with robots.

3.2.2 Target Population

The target population for this study was exclusively adults from the general population aged 18 or over with normal to corrected to normal vision, no known mental disability, hearing loss or difficulties, or physical handicap, native English speaking, and currently residing in Great Britain. Due to the technical requirements of the mediated experimental design, the target population of this study consisted of individuals with access to a personal computer with Zoom installed, a functioning web camera, a stable internet connection, a microphone, and speakers/headphones.

3.2.3 Recruitment

Participants were recruited via Prolific and were allowed to participate only after confirming that they were older than 18 years, are native English speakers, and have access to a computer with Zoom installed as well as a decent web camera, stable internet connection, microphone, and speakers/headphones. Also, Prolific users were asked to commit to attending 2 sessions a week across 5 weeks. Eligible Prolific users could access the Prolific page of the study to receive further information, consider their participation, and complete the induction questionnaire if interested. On the Prolific page of the study (of the induction questionnaire—Session 0) and in the induction questionnaire Qualtrics form, Prolific users were introduced to the study, the task, and the available time slots as part of the longitudinal experiment schedule. After receiving this information about the study's requirements, Prolific users were then asked if they would like to continue in the study by declaring that they can commit to the study's requirements. Finally, Prolific users were then asked to choose their participation time slots, after which they received a participant number to start their participation. Participants were paid a total of £3 for every 30 min of participation or participation session if it lasted less than 30 min. Participants who completed all 10 sessions were paid an extra £20 after their final interaction. A detailed description of the recruitment procedure and a full list with specific Prolific filters used for participant recruitment can be found in the study's OSF page [see 86].

3.2.4 Ethics and Communication

All study procedures were approved by the research ethics committee of the University of Glasgow (ethics approval numbers 300200094 & 300200132). All participants provided written informed consent before participating in the study. Participants were asked to provide, if they wished, optional consent to allow the research team to use their video and audio footage (including videos, audio, and photos made from video material) as materials for research publications, conference presentations, and other multimedia outputs that can and might be disseminated and distributed online, in the media and for public presentations. All Prolific users interested in participating in the study were introduced to the study, the requirements of the study, and the task, but were not informed about the functionalities of the robot Pepper, to ensure all robot knowledge or priming was minimised. During each session (including Session 0), participants were re-introduced to the study, the study's schedule (about their chosen day of participation), and received reminders and information about what the study involves. Furthermore, they were reminded about the benefits and risks of their participation (i.e., ensuring that they would receive their payment, no risks were anticipated as a result of study involvement, and their right to withdraw their participation at any time with no penalty or punishment). Participants were further informed how their data (i.e., behavioural and self-reported data collected in the study) would be used and again reminded of their right to withdraw their data and/or ask that it not be used at any time during or after their participation. Participants were guaranteed that their right to privacy and anonymity would be respected and that no identifiable data would be shared with anyone beyond the research team. Participants were reminded that their participation was voluntary and they were given the contact information of the main researcher and experimenter should they wish to follow up with any further questions. After completing the study, participants received a comprehensive debriefing message in Prolific (forwarded by Prolific to their associated email address), providing further information about the study, the deception that was used (i.e., the experimenter was using WoZ approach for communicating with participants to make it look like the robot was responding autonomously), and were again given the contact information of the main researcher and experimenter should they wish to follow up with any further questions or feedback.

When completing each session, participants were reminded in the Qualtrics form about the date of their next session. Two days before each session, participants received an email via Prolific regarding the specifics of their next session. This message contained details about the session number, the time at which the Prolific page with the link for the session questionnaire form would be published, a reminder not to start the session before the allocated participation time slot, and to contact the experimenter if they are to be late, cannot remember their participation time-slot, or cannot make it. Finally, participants were thanked for their participation and cooperation in each of these messages and were reminded of their rights and the fact that they were welcome to contact the experimenter at any time using the Prolific messaging system, or by email. On the day of participation, participants received an automated message from Prolific at 08:00 AM, that the Prolific page of the session is available online. Later that day, 30 to 15 min before each participant's participation time slot, each participant received an individual message via Prolific from the experimenter about their upcoming session and where they could find the link to start the session. If and when participants were late to their participation (without providing earlier notice that this would be the case), the experimenter messaged the participant via Prolific to ensure attendance or reschedule the session. When participants experienced any technical difficulties or needed to communicate with the experimenter, they were instructed to do this via Prolific or email, and not using the Zoom chat. This was to reduce any potential association between the session interactions with the robot and the experimenter. Accordingly, all communications between participants and the experimenter took place via the Prolific messaging centre or emails on rare occasions (when initiated by a participant). The main researcher and experimenter (GL) signed his name on all communications with participants.

3.3 Stimuli

Conversational interactions were guided by the robot Pepper (SoftBank Robotics), a humanoid robot capable of communicating via speech and gestures. Following [26] guidelines for social robots' design for long-term interactions, Pepper was chosen as a suitable robotic platform for this task, given the alignment between Pepper's humanoid embodiment and the social requirements of the conversa-

Fig. 1 The lab settings, including the robot Pepper (SoftBank Robotics) in front of a web camera, while the experimenter in the back is controlling the robot using the Wizard of Oz technique
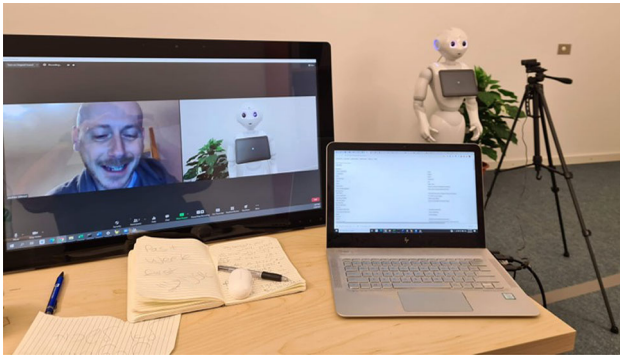
Fig. 2 The interaction from the eyes of the participants and the experimenter. The participants were exposed only to the robot Pepper (SoftBank Robotics) via the zoom chats

tional task [see 26; "Guidelines for Future Design"]. While Pepper's appearance and behaviours are somewhat human-like (i.e., Pepper has a head, face, torso, two arms, two hands, five fingers per hand, etc.), Pepper has not been designed to resemble a real person. Instead, Pepper's embodiment and behaviours clearly convey human likeness, (further evidenced, for example, by Pepper's abilities to communicate using human speech, but not demonstrating any facial expressions given the rigid, immobile face and head).

Pepper was placed in front of a web camera (Logitech, 1080p), connected to the experimenter's computer (see Fig. 1). Behind Pepper was a white wall and a flowerpot with a green plant (see Fig. 2). Pepper communicated with participants in this study via the WoZ technique controlled by the experimenter via a PC laptop. All pre-scripted questions and speech items were written and coded in the WoZ system, with the experimenter controlling Pepper by pressing buttons on a PC laptop. Accordingly, the procedure followed a clear pre-programmed protocol where the experimenter did not need to speak or type anything during the interaction, but only pressed the relevant keys to trigger the required or appropriate text delivery via Pepper.

Pepper responded to participants' answers and statements with neutral or empathetic responses. Pepper's vocabulary was limited and constrained to reflect the current state of speech recognition technology in social robotics. Following [26] guidelines for social robots' design for long-term interactions, Pepper's responses were affective and empathetic, aiming to convey an understanding of users' affective state, communicate appropriate responses, and also display contextualised affective reactions [see 26; "Guidelines for Future Design"]. Hence, a limited set of responses were pre-defined for answers and statements with neutral sentiment or containing factual information (e.g., "I understand", "I see", "okay"), for answers and statements of positive sentiment (e.g., "I am happy to hear that", "This is really interesting", "That's amazing"), and for answers and statements of nega-

tive sentiment (e.g., "I am sorry to hear that", "This sounds very challenging", "These are not easy times"). Moreover, Pepper had pre-defined statements for opening an interaction (e.g., "Hello there", "Hi!", "How are you doing today?"), closing an interaction ("That's it for now", "See you next time", "Have a good weekend", "Goodbye"), answer with basic polite gratitude (e.g., "I am fine, thank you!", "Thank you", "That is lovely of you to say so", "It was nice to chat with you too!"), and thank participants for their cooperation and disclosures (e.g., "Thank you for sharing with me", "Thank you for telling me", "What a nice memory. Thank you for sharing with me"). Due to Pepper's high-pitched voice and robotic style of pronunciation, Pepper's answers and statements were structured using commas so that Pepper's speech segments will be clearer. See the OSF repository [86] for a file with all of Pepper's vocabulary and the structure of Pepper's speech segments.

Pepper communicated using a cheerful, high-pitched voice, and expressive and animated body language that corresponded to the spoken content and Pepper's physical capabilities. Pepper's movements were self-initiated based on Pepper's demo software's "Animated Speech" function[1], in order to provide a sense of neutral interaction and to ensure replicability by future studies using the same functionality that all Pepper robots are equipped with. Moreover, Pepper's gaze was almost always focused on the camera, but it shifted and moved from the camera with no pre-programmed logic by leaving the "Basic Awareness" function[2] on. To ensure that the mediated interactions would come across as natural, Pepper's gaze was not programmed to be focused on the camera at all times as this would not be normal behaviour with conversing with a human interlocutor. Therefore, Pepper's gaze shifts were allowed to naturally occur following its demo software.

3.4 Behavioural Paradigm

The behavioural paradigm was specifically designed to encourage participants to engage in self-disclosure behaviour. By asking them questions about personal matters, we aimed to create an environment conducive to self-disclosure during interactions with the social robot. The intention behind this design was to facilitate meaningful exchanges and encourage participants to share personal information with the robot. In accordance with [26] guidelines for social robots' design for long-term interactions, the interactions followed a clear structure and routine, including greetings and farewells, identifying participants by their name, and demonstrating appropriate affective and emphatic responses to participants'

---

[1] http://doc.aldebaran.com/2-5/naoqi/audio/alanimatedspeech.html.

[2] https://qisdk.softbankrobotics.com/sdk/doc/pepper-sdk/ch4_api/abilities/reference/BasicAwareness.html.

answers to provide a sense of personal interactions and encourage self-disclosure [see 26; "Guidelines for Future Design"].

3.4.1 Structure

Each interaction was guided by Pepper as a semi-structured interview discussing non-sensitive topics regarding general everyday experiences. Each interaction followed the same order, starting with greetings followed by 3 questions (x3 repetitions). The participants were instructed to have a short conversation with Pepper, following Pepper's lead in the interaction and answering Pepper's questions. Participants were instructed that no time limit was applied for the interactions and that the interactions usually took about five to ten minutes. They were further encouraged to participate in the interactions the way they saw fit—speaking as little or as much as they wished. In addition, participants were instructed that there were no correct or incorrect answers, and they were encouraged to provide honest answers according to what they felt comfortable with. In the first interaction with Pepper (Session 1), participants were asked for their name by the robot as part of the robot introduction (i.e., "Hello there, my name is Pepper, what is your name?"), as such a question would be part of a normal introduction in on-going social exchanges with another person. Before the interaction started, participants were instructed that they were not obliged to share their names with the robot and that they could give a fake name if they preferred to do so. From the second interaction (Session 2) onwards, Pepper addressed each participant by the name they gave during the first interaction (Session 1), to provide a sense of natural and personalized interactions. The task followed the following structure and order:

- Short greetings/introduction (e.g., Hi there, how are you doing?).
- One pre-defined general question about the participant's day, week, or weekend, to build rapport (e.g., "how was your weekend? Did you do anything interesting?").
- An opening statement introducing the topic of the question (e.g., "I am about to ask you about your social life").
- Two pre-defined, non-sensitive questions that correspond to the topic that was randomly allocated to the interaction. These questions were either framed in the context of the COVID-19 pandemic or in a more general everyday context, depending on the discussion frame group assignment.

3.4.2 Content

Previous studies that investigated relationship formation and disclosure with artificial agents followed conceptual frameworks for inducing rich disclosures and forming meaningful connections [e.g., 3, 52, 53, 89]. For example, a study by [53] presented an implementation of 36 questions as a method to generate interpersonal closeness [see 90; "36 questions to love"] and elicit self-disclosure from human users to a chatbot. A previous study [3] demonstrated how simple questions about everyday experiences (i.e., work-life balance and finances, social life and relationships, and health and well-being) can elicit meaningful disclosures when communicated by a social robot. The questions and topics in the study were influenced by [12] and [91] as an elicitation technique aiming to capture participants' subjective experiences regarding various everyday topics. Here we used a similar type of questions to the ones used in [3], adapting disclosure topics for the ten sessions from [91] and [92], and framing the disclosure topics and questions in this study following a framework and guidelines by [92].

The framework by [92] introduces guidelines via six main themes for asking questions that capture and elicit disclosures that relate to different elements of quality of life within counselling psychology settings and mental health therapy. The guidelines and themes were defined by [92] after reviewing and synthesizing qualitative research studies (especially from the counselling psychology literature, psychotherapy, and mental health therapy literature) that explicitly asked adult participants with mental health problems about the factors they considered important to their quality of life or how it had been impacted by their mental health. Based on [92] review's results the six themes are: (1) Well-being and Ill-being, (2) Control, Autonomy, and Choice, (3) Self-Perception, (4) Belonging, (5) Activity, (6) Hope and Hopelessness.

The ten topics for the ten sessions describe one or more of the six themes described by [92], aiming to elicit meaningful disclosures following [91] guidelines, but also to initiate self-reflection and capture meaningful information regarding the quality of life and mental health, following

Table 1 The ten topics and corresponding quality of life categories following [92] framework

	Topic	Category
1	Work	Control, Autonomy, and Choice
2	Leisure and Passions (Life)	All
3	Finances	Control, Autonomy, and Choice
4	Relationships	Belonging
5	Social life	Belonging
6	Mental Health	Well-being and Ill-being
7	Physical Health	Well-being and Ill-being
8	Personality	Self-Perception
9	Goals and Ambitions	Hope and Hopelessness
10	Routine and Daily Activities	Activity

[92] framework. The ten topics and their corresponding themes according to [92] framework can be seen in Table 1. The phrasing of each of the two questions under each topic followed [90] approach for questions and practical methodology for creating interpersonal closeness in an experimental context (see the questions in the OSF repository [86]).

3.4.3 Discussion Frames

For both discussion frames, the interaction always started the same way, with greetings and with the robot asking the first question about the participant's day/week/weekend (see Sect. 3.4.1 for the structure of the task). The following two pre-defined questions were about a topic that was randomly allocated to the interaction from the 10 topics about general everyday experiences (see Table 1). For participants assigned to the neutral discussion frame group, the questions were not limited to any specific frame other than general everyday context. For participants assigned to the Covid-related discussion frame group, questions were asked about the same topics, however, the questions were framed within the context of the COVID-19 pandemic. For example, participants were asked how their work situation changed due to the pandemic, or how they were socializing during the pandemic. See the questions and differences between conditions at the OSF repository [86].

3.5 Measurements

To ensure that our models only include high-quality data, we included only cases that were captured and processed correctly.

3.5.1 Demographics

Participants were requested to complete a short questionnaire that gathered information on demographic parameters including age, biological sex, gender identification, level of education, nationality, job, previous experience with robots, and whether English is their native language.

3.5.2 Self-Disclosure

We operationalized the concept of self disclosure via three angles. Single dimensions cannot capture the complex nature of self-disclosure, as it is a multidimensional behaviour [93]; perceptions of self-disclosure can be subjectively reported and objectively observed differently from behaviour and content [94]. First, we measured participants' subjective perceptions of their own self disclosures towards Pepper to gain subjective insights about the interaction from the participants point of view. This method is practical as it is easy to administer and provides a subjective perspective on self-disclosure. However, it may be subject to bias, as participants may not accurately report their behaviour. To capture self-disclosure behaviour more objectively, we focused on two additional, quantitative, measures: the duration of all utterances in an answer (i.e., self disclosure duration in seconds) and the average word count per answer (i.e., self disclosure length in number of words). These measures allowed us to examine the volume of disclosure during the interactions. Self-disclosure has been linked to the total number of words a person produces during an interaction or within a single turn in the interaction. Higher word counts are associated with greater self-disclosure [95–97]. Therefore, these measures provided further objective criteria to assess changes in disclosure behaviour and its relationship with the social robot. Finally, to capture the emotional and sentimental tone of the disclosure we used a measure of disclosure compound sentiment, rating the sentiment present in the disclosure from negative to positive.

Subjective self-disclosure
Participants were requested to report their level of subjective self-disclosure via the sub-scale of work and studies disclosure in Jourard's Self-Disclosure Questionnaire [91]. This questionnaire was adapted and adjusted for the context of the study, addressing the statements to general life experiences. The measurement included ten self-reported items for which participants reported the extent to which they disclosed information to Pepper on a scale of one (not at all) to seven (to a great extent). Accordingly, a mean scale was constructed ($M = 3.60$, $SD = 1.17$) which was found to be reliable (Cronbach's $\alpha = .83$).

Disclosure duration
Duration of speech in seconds from each recording was extracted and processed using Parselmouth [98], a Python library for Praat [99].

Disclosure length
The volume of disclosure in terms of the number of words per disclosure. The recordings were automatically processed using the IBM Watson speech recognition engine, applying the British telephony model. To ensure capturing all utterances within each disclosure we amplified the audio files with 7 decibels and slowed the audio file's pitch by adjusting the sample rate, with a crucial factor of 0.55. By reducing the sample rate using the key factor of 0.55, we effectively decreased the playback speed of the recordings. This method was used only for receiving text output for analysis and not for the robotic interactions. This method was used systematically across all data units in the study to ensure that the procedure would be identical across all data units. The number of words per disclosure was extracted from the text using a simple length command in Python.

Disclosure compound sentiment

Using Vader for Python [100], the disclosures were measured to determine their overall sentiment in terms of positive, neutral, and negative sentiment. The compound sentiment evaluates a disclosure sentiment from negative (-1) to positive (+1), based on the calculated sentiment score [see 100].

3.5.3 Perception

Agency and experience

Research into mind perception has revealed that agency (the ability of an agent to plan and act) and experience (the ability of the agent to sense and feel) are two key dimensions when valuing an agent's mind [101]. To determine whether any differences in mind perception emerged across the testing sessions, participants were requested to evaluate Pepper in terms of agency and experience, after being introduced to these terms [adapted from 101]. Both concepts were evaluated by the participants using a 0 to 100 rating bar.

Friendliness and warmth

This scale was aimed at capturing how participants perceived Pepper in terms of friendliness and warmth using one item from [102] and two items from [103], as suggested by [54]. These items were evaluated on a seven-point scale ranging from 1 (not at all) to 7 (extremely). Accordingly, a mean scale was constructed ($M = 6.11$, $SD = 1.02$) which was found to be reliable (Cronbach's α =.94).

Communication competency

This scale was aimed at capturing how participants experienced and evaluated Pepper's communication competency using an adapted and adjusted version by [52] for a scale by [104]. The scale included three items that were evaluated on a seven-point scale ranging from 1 (not at all) to 7 (extremely). Accordingly, a mean scale was constructed ($M = 5.78$, $SD = 1.18$) which was found to be reliable (Cronbach's α =.93).

Interaction quality

This scale was aimed at capturing how participants perceived and evaluated the interaction with Pepper using an adapted and adjusted version by [52] for a scale by [105]. Each interaction included two random items out of seven, except for the mid-session (session 5) and the last session (session 10) which included all six items of the scale. These items were evaluated on a seven-point scale ranging from 1 (not at all) to 7 (extremely). Accordingly, a mean scale was constructed ($M = 5.48$, $SD = 1.56$) which was found to be reliable (Cronbach's α =.96).

3.5.4 Well Being

Mood

To capture participants' mood change from their interactions with Pepper, participants reported their mood before and after the interaction with Pepper using the Immediate Mood Scaler [IMS-12; see 106]. IMS-12 includes 12 items of polarized moods, ranging from 1 (for negative moods) to 7 (for the equivalent positive moods). The scale is a novel validated tool based on the Positive and Negative Affect Schedule [PANAS; 107], adapted and adjusted to capture current mood states in online and mobile experiments [106]. Mean reliable scales were constructed for participants' mood before the interaction ($M = 5.35$, $SD = 1.16$, Cronbach's α =.96) and after the interaction ($M = 5.75$, $SD = 1.08$, Cronbach's α =.97).

Comforting responses

To measure the extent to which participants perceived Pepper's responses as comforting the comforting response scale was adapted [see 108]. The scale includes 12 self-reported items rated on a seven-point scale, ranging from 1 (I strongly disagree) to 7 (I strongly agree). Accordingly, a mean scale was constructed ($M = 5.50$, SD =.89) which was found to be reliable (Cronbach's α =.91).

Loneliness

Each session participants were requested to report their feelings and thoughts of loneliness from the last three days using the short-form UCLA loneliness scale [ULS-8; see 109]. The scale includes 8 items rated on a seven-point scale, ranging from 1 (not at all) to 7 (all the time). Accordingly, a mean scale was constructed ($M = 2.86$, $SD = 1.28$) which was found to be reliable (Cronbach's α =.90).

Stress

Participants were requested to report their feelings and thoughts of periodic stress from the past month using the perceived stress scale [110]. The scale includes 10 statement items rated on a seven-point scale, ranging from 1 (never) to five (very often). A mean scale was constructed ($M = 3.30$, $SD = 1.03$) which was found to be reliable (Cronbach's α =.89).

3.6 Materials

3.6.1 Zoom Video Chat

All interactions (video chats) were conducted with the software Zoom, using a university staff account (see Fig. 2). The interactions were recorded using the recording functionality on Zoom and edited to include only those portions of the recordings where participants and/or Pepper were speaking.

3.6.2 Qualtrics Questionnaires

All of the questionnaires were administered via the survey software Qualtrics, using a university staff account. In the online questionnaires, the functionality of recording partic-

ipants' IP addresses was disabled to comply with GDPR guidelines.

3.7 Procedure

When recruited, participants completed an induction questionnaire (Session 0) approximately one week before beginning their video chat interactions with Pepper (Sessions 1 to 10). Participants were instructed to have a short conversation with Pepper about several topics that Pepper will bring up, that Pepper will ask them 3 questions and that the interactions will take place twice a week across five weeks during prearranged times. They were further told that each interaction with Pepper should last about 5 to 10 min, and another 10–15 min will be required to complete questionnaires afterwards. When answering the induction questionnaire (after providing consent to participate in the study), participants were instructed on how to position their video camera for the video chats, and what the lighting in the room is expected to be like. Following this, participants reported on several demographic parameters and several questionnaires (for the full list of questionnaires and their order in each session see the OSF repository at [86]). Participants were redirected to the Prolific website when completing the induction questionnaire (Session 0). A participant number was automatically generated for each participant who completed the induction questionnaire (Session 0) and proceeded to the following sessions. The random assignment of participants to conditions, allocation of topics to sessions for each participant and the order of questions in each interaction were randomized and allocated automatically and an excel sheet was created to help the experimenter control and follow the experimental design procedure for five weeks (see the randomization and allocation code, experimenter notebook with the conditions, allocated topics to a session, and order of questions for each of the participants on the OSF repository at [86]).

When starting each session, participants were asked to enter their Prolific ID and their participant number. Following, participants were asked to answer the Immediate Mood Scale [IMS-12, 106] for reporting their mood before interacting with Pepper. Next, participants received a reminder regarding their interaction with Pepper, what the task requires, and some basic instructions. The page included a link to the Zoom interaction, a frame with the zoom landing page, and the experimenter's e-mail address and instructions on how to communicate with the experimenter in case there are any issues during the interaction. Then, participants interacted with Pepper via a Zoom video chat (see Sect. 3.4), only seeing Pepper in the chat (see Fig. 2). After finishing their interaction with Pepper, participants went back to the Qualtrics page and answered the rest of the questionnaires. The full list of questionnaires and their order in each session

can be found on the study's OSF page [see 86]. When finished answering the questionnaire, participants were thanked for the completion of the session, reminded about the date and day of their upcoming session, were provided again with the contact details of the experimenter, and were directed back to Prolific to receive a completion message. When completing the last session participants were clarified that this is indeed the last session, they were thanked for their participation, and provided with contact details of the experimenter to ask any further questions about the study.

4 Results

4.1 Disclosure

We used lme4 [111] for R to perform a linear mixed effects analysis of the effect of session number, discussion frame and their interaction term on participants' disclosures to Pepper. As fixed effects, we entered the session order, the discussion frame and their interaction term into the model. As a random effect, we had intercepts for subjects. Significance was calculated using the lmerTest package [112], which applies Satterthwaite's method to estimate degrees of freedom and generate p-values for mixed models.

4.1.1 Subjective Self-Disclosure

The model explains 60.5% of the variance in participants' *subjective perceptions of their self-disclosure* to Pepper, whereas the fixed effects in the model explain 3.4% of the variance (see Table 2). The results stress that despite the variance between the participants (SD =.90), the session number has a significant positive fixed effect on participants' subjective perceptions of their self-disclosures (β =.07, SE =.02, p < .001, see Fig. 3). Nevertheless, there were no significant fixed effects in terms of the discussion frame ($\beta = -.16, SE = .33, p = .627$), and the interaction term of the session number and discussion frame ($\beta = -.02, SE = .03, p = .529$).
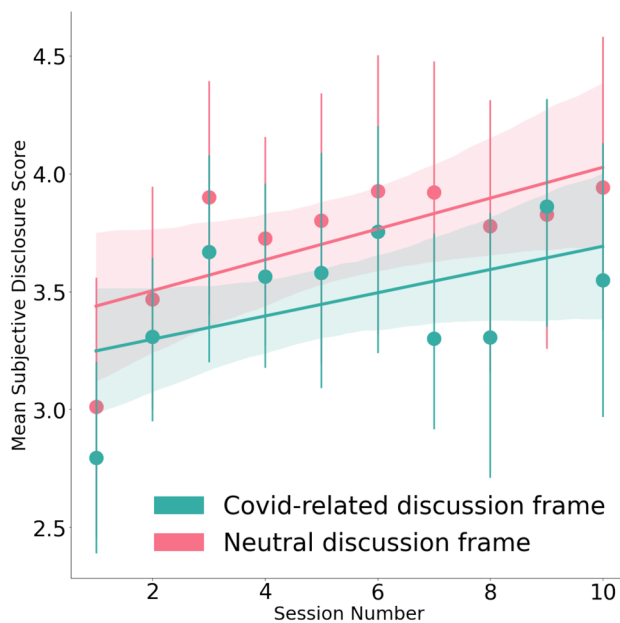
4.1.2 Disclosure Duration

The model explains 49.1% of the variance in participants' *disclosures duration* (in seconds) to Pepper, whereas the fixed effects in the model explain 3.7% of the variance (see Table 2). The results stress that despite the variance between the participants ($SD = 21.04$), the session number has a significant positive fixed effect on participants' disclosures duration ($\beta = 2.10, SE$ =.33, p < .001, see Fig. 4). Nevertheless, there were no significant fixed effects in terms of the discussion frame ($\beta = 4.04, SE = 7.32, p$ =.583), and the

Table 2 Results of disclosure

Fixed Effects	Subjective Disclosure		Duration		Length		Compound Sentiment	
	Estimates	95%CI	Estimates	95%CI	Estimates	95%CI	Estimates	95%CI
Intercept	3.33***	2.87 − 3.80	17.62***	7.34 − 27.90	40.78***	16.71 − 64.86	0.41***	0.31 − 0.52
Discussion frame	-0.16	-0.82 − 0.49	4.04	-10.31 − 18.39	9.11	-24.51 − 42.72	-0.04	-0.19 − 0.10
Session number	0.07***	0.03 − 0.11	2.09***	1.45 − 2.74	4.97***	3.47 − 6.47	0.02***	0.01 − 0.04
Discussion frame * Session number	-0.02	-0.07 − 0.04	-0.13	-1.03 − 0.77	-0.09	-2.18 − 2.01	-0.01	-0.02 − 0.01
Random Effects								
SD	0.90		21.04		49.32		0.15	
σ^2	0.56		497.29		2692.53		0.22	
τ_{00}	0.81		442.54		2432.56		0.02	
ICC	0.59		0.47		0.47		0.09	
N	39		39		39		39	
Observations	386		1160		1160		1160	
Marginal R^2 / Conditional R^2	0.034 / 0.605		0.037 / 0.491		0.041 / 0.496		0.021 / 0.108	

*Note: * p<0.05 ** p<0.01 *** p<0.001*

Fig. 3 Mean subjective disclosure scores by session number and discussion frame

interaction term of the session number and discussion frame ($\beta = -.13$, $SE = .46$, $p = .774$).

Another linear mixed effects model was used to test if the discussion frame, the session number, and their interaction term significantly predicted the *disclosure duration* when interacting with the social robot Pepper, including only the items corresponding to the disclosure topic. The model explains 61.1% of the variance in participants' disclosures duration (in seconds) to Pepper, whereas the fixed effects in the model explain 5.1% of the variance (see Table 3). The results reveal that despite the variance between participants

($SD = 26.53$), the session number has a significant positive fixed effect on participants' disclosures duration ($\beta = 2.54$, $SE = .40$, $p < .001$, see Fig. 4). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = 6.50$, $SE = 9.18$, $p = .482$), and the interaction term of the session number and discussion frame ($\beta = .03$, $SE = .56$, $p = .964$).

4.1.3 Disclosure Length

The model explains 49.6% of the variance in participants' *disclosures length* (in number of words) to Pepper, whereas the fixed effects in the model explain 4.1% of the variance (see Table 2). The results stress that despite the variance between the participants ($SD = 49.32$), the session number has a significant positive fixed effect on participants' disclosures length ($\beta = 4.97$, $SE = .76$, $p < .001$, see Fig. 5). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = 9.11$, $SE = 17.13$, $p = .598$), and the interaction term of the session number and discussion frame ($\beta = -.09$, $SE = 1.07$, $p = .936$).

Another linear mixed effects model was used to test if the discussion frame, the session number, and their interaction term significantly predicted the *disclosure length* when interacting with the social robot Pepper, including only the items corresponding to the disclosure topic. The model explains 61.6% of the variance in participants' disclosures length (in number of words) to Pepper, whereas the fixed effects in the model explain 5.6% of the variance (see Table 3). The results stress that despite the variance between the participants ($SD = 61.89$), the session number has a significant positive fixed effect on participants' disclosures length ($\beta = 6.09$, $SE = .92$, $p < .001$, see Fig. 5). Nevertheless, no significant fixed effects emerged in terms of the discussion frame
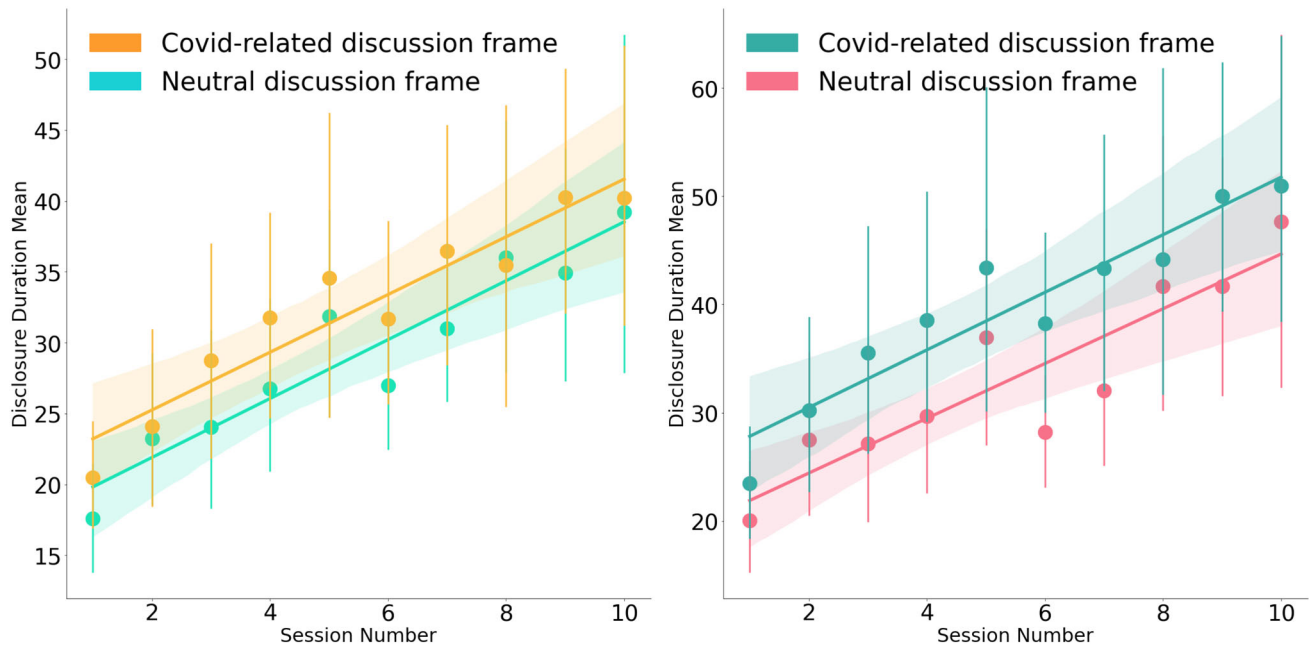
Fig. 4 From left to right: (1) Mean disclosure duration (in seconds) by session number and discussion frame. (2) Mean disclosure duration (in seconds) by session number and discussion frame, including only the items corresponding to the disclosure topic

Table 3 Results of disclosure including only the items that corresponded to the topic of disclosure

	Duration		Length		Compound Sentiment	
Fixed Effects	*Estimates*	*95%CI*	*Estimates*	*95%CI*	*Estimates*	*95%CI*
Intercept	19.29**	$6.39 - 32.19$	44.20**	$14.14 - 74.26$	0.48***	$0.36 - 0.59$
Discussion frame	6.50	$-11.51 - 24.51$	15.58	$-26.40 - 57.56$	-0.07	$-0.24 - 0.10$
Session number	2.54***	$1.76 - 3.32$	6.09***	$4.28 - 7.90$	0.02**	$0.01 - 0.04$
Discussion frame * Session number	0.02	$-1.07 - 1.12$	0.24	$-2.29 - 2.78$	-0.01	$-0.03 - 0.02$
Random Effects						
SD		26.53		61.89		0.13
σ^2		488.40		2626.54		0.22
τ_{00}		703.83		3830.41		0.02
ICC		0.59		0.59		0.07
N		39		39		39
Observations		773		773		773
Marginal R^2 / Conditional R^2		0.051 / 0.611		0.056 / 0.616		0.026 / 0.094

*Note: * p<0.05 ** p<0.01 *** p<0.001*

($\beta = 15.58$, $SE = 21.39$, $p = .470$), and the interaction term of the session number and discussion frame ($\beta = .24$, $SE = 1.29$, $p = .852$).

4.1.4 Disclosure Compound Sentiment

The model explains 10.8% of the variance in participants' disclosures *compound sentiment* (see Sect. 3.5.2), whereas the fixed effects in the model explain 2.1% of the variance

(see Table 2). The results stress that despite the variance between the participants ($SD = .15$), the session number has a significant positive fixed effect on participants' disclosures compound sentiment ($\beta = .02$, $SE = .01$, $p < .001$). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -.04$, $SE = .08$, $p = .569$), and the interaction term of the session number and discussion frame ($\beta = -.01$, $SE = .01$, $p = .537$).
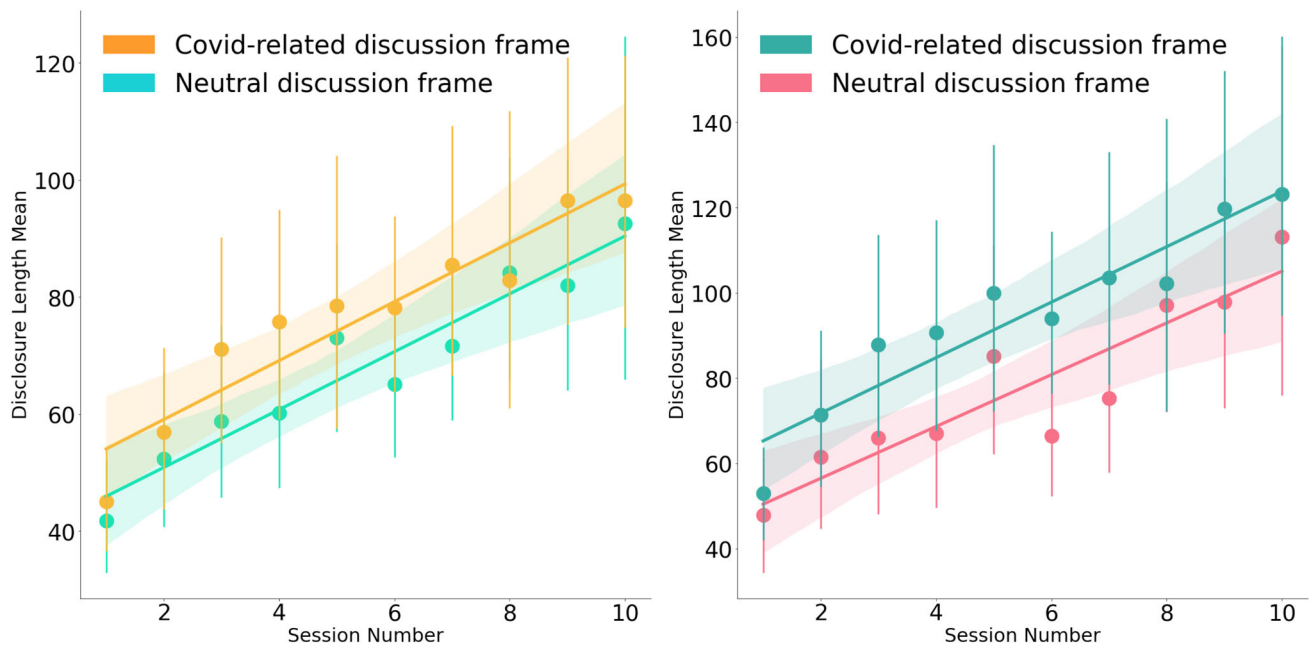
Fig. 5 From left to right: (1) Mean disclosure length (in number of words) by session number and discussion frame. (2) Mean disclosure length (in number of words) by session number and discussion frame, including only the items corresponding to the disclosure topic

Another linear mixed effects model was run to test if the discussion frame, the session number, and their interaction term significantly predicted the disclosure *compound sentiment* when interacting with the social robot Pepper, including only the items that corresponded to the topic of disclosure. The model explains 9.4% of the variance in participants' disclosures compound sentiment to Pepper, whereas the fixed effects in the model explain 2.6% of the variance (see Table 3). The results stress that despite the variance between the participants (SD =.13), the session number has a significant positive fixed effect on participants' disclosures duration (β =.02, SE =.01, p = .005). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -.07$, SE =.09, p =.418), and the interaction term of the session number and discussion frame ($\beta = -.01$, SE =.01, p =.575).

4.2 Perception

We used lme4 [111] for R to perform linear mixed effects analysis of the effect of session number, discussion frame and their interaction term on participants' perceptions of Pepper, including perceptions of agency and experience [see 101], friendliness and warmth, communication competency and interaction quality. As fixed effects, we entered the session order, the discussion frame and their interaction term into the model. As a random effect, we had intercepts for subjects. Significance was calculated using the lmerTest package [112], which applies Satterthwaite's method to esti-

mate degrees of freedom and generate p-values for mixed models.

4.2.1 Agency

The model explains 82.5% of the variance in participants' perceptions of Pepper's degree of *agency*, whereas the fixed effects in the model explain 1.6% of the variance (see Table 4). The results stress that despite the variance between the participants ($SD = 21.38$), the session number has a significant positive fixed effect on participants' perceptions of Pepper's degree of agency ($\beta = 1$, SE =.25, $p < .001$, see Fig. 6). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -1.34$, $SE = 7.20$, p =.853), and the interaction term of the session number and discussion frame (β =.08, SE =.35, p =.828).

4.2.2 Experience

The model explains 79.4% of the variance in participants' perceptions of Pepper's degree of *experience*, whereas the fixed effects in the model explain 3.8% of the variance (see Table 4). The results stress that despite the variance between the participants ($SD = 24.29$), the session number has a significant positive fixed effect on participants' perceptions of Pepper's degree of experience ($\beta = 1.82$, SE =.32, $p < .001$, see Fig. 6). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -3.59$, $SE = 8.27$, p

Table 4 Results of perception

Fixed Effects	Agency		Experience		Friendliness and Warmth		Communication Competency		Interaction Quality	
	Estimates	95%CI	Estimates	95%CI	Estimates	95%CI	Estimates	95%CI	Estimates	95%CI
Intercept	62.61***	52.48 − 72.74	55.66***	44.01 − 67.30	5.83***	5.38 − 6.27	5.61***	5.12 − 6.11	4.97***	4.33 − 5.60
Discussion frame	-1.34	-15.49 − 12.81	-3.59	-19.386 − 12.67	-0.07	-0.69 − 0.55	-0.07	-0.85 − 0.53	-0.21	-1.10 − 0.68
Session number	1***	0.50 − 1.49	1.82***	1.18 − 2.45	0.05***	0.02 − 0.07	0.03*	0.00 − 0.07	0.09***	0.05 − 0.14
Discussion frame * Session number	0.08	-0.62 − 0.77	-0.11	-1.00 − 0.78	0.02	-0.01 − 0.05	0.02	-0.03 − 0.06	0.03	-0.03 − 0.10
Random Effects										
SD	21.38		24.29		0.93		1		1.26	
σ^2	98.61		160.95		0.23		0.43		0.85	
τ_{00}	457.23		589.89		0.86		1		1.59	
ICC	0.82		0.79		0.79		0.70		0.65	
N	39		39		39		39		39	
Observations	386		386		386		386		386	
Marginal R^2 / Conditional R^2	0.016 / 0.825		0.038 / 0.794		0.025 / 0.797		0.012 / 0.703		0.041 / 0.664	

*Note: * $p<0.05$ ** $p<0.01$ *** $p<0.001$*

=.666), and the interaction term of the session number and discussion frame ($\beta = -.11$, SE =.45, p =.811).

4.2.3 Friendliness and Warmth

The model explains 79.7% of the variance in participants' perceptions of Pepper's degree of *friendliness and warmth*, whereas the fixed effects in the model explain 2.5% of the variance (see Table 4). The results stress that despite the variance between the participants (SD =.93), the session number has a significant positive fixed effect on participants' perceptions of Pepper's degree of friendliness and warmth (β =.05, SE =.01, $p < .001$). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -.07$, SE =.32, p =.828), and the interaction term of the session number and discussion frame (β =.02, SE =.02, p =.245).

4.2.4 Communication Competence

The model explains 70.3% of the variance in participants' perceptions of Pepper's *communication competence*, whereas the fixed effects in the model explain 1.2% of the variance (see Table 4). The results stress that despite the variance between the participants ($SD = 1$), the session number has a significant positive fixed effect on participants' perceptions of Pepper's communication competence (β =.03, SE =.02, $p = .040$). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -.16$, SE =.35, p =.655), and the interaction term of the session number and discussion frame (β =.02, SE =.02, p =.456).

4.2.5 Interaction Quality

The model explains 66.4% of the variance in participants' perceptions of the *interaction quality*, whereas the fixed effects in the model explain 4.1% of the variance (see Table 4). The results stress that despite the variance between the participants ($SD = 1.26$), the session number has a significant

positive fixed effect on participants' perceptions of the interaction quality (β =.09, SE =.02, $p < .001$). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -.21$, SE =.45, p =.646), and the interaction term of the session number and discussion frame (β =.04, SE =.03, p =.291).

4.3 Well-Being

We used lme4 [111] for R to perform linear mixed effects analysis of the effects of session number, discussion frame and their interaction term on participants' perceptions of Pepper's comforting responses, mood change, and feelings of loneliness. As fixed effects, we entered the session order, the discussion frame and their interaction term into the model. As a random effect, we had intercepts for subjects. Significance was calculated using the lmerTest package [112], which applies Satterthwaite's method to estimate degrees of freedom and generate p-values for mixed models.

4.3.1 Mood

The model explains 69.8% of the variance in participants' *mood*, whereas the fixed effects in the model explain 4.2% of the variance (see Table 5). The results stress that despite the variance between the participants (SD =.94), we observed a positive significant fixed effect on mood change, as participants reported a positive mood change after interacting with Pepper (β =.49, SE =.11, $p < .001$). Moreover, the session number has a significant positive fixed effect on participants' mood (β =.03, SE =.33, $p = .019$, see Fig. 7). Nevertheless, no significant fixed effects emerged in terms of the discussion frame ($\beta = -.24$, SE =.32, p =.469), the interaction term of the session number and discussion frame (β =.02, SE =.02, p =.154), the interaction term of the session number and mood change ($\beta = -.01$, SE =.02, p =.388), and the interaction term of the discussion frame and mood change ($\beta = -.01$, SE =.09, p =.943).
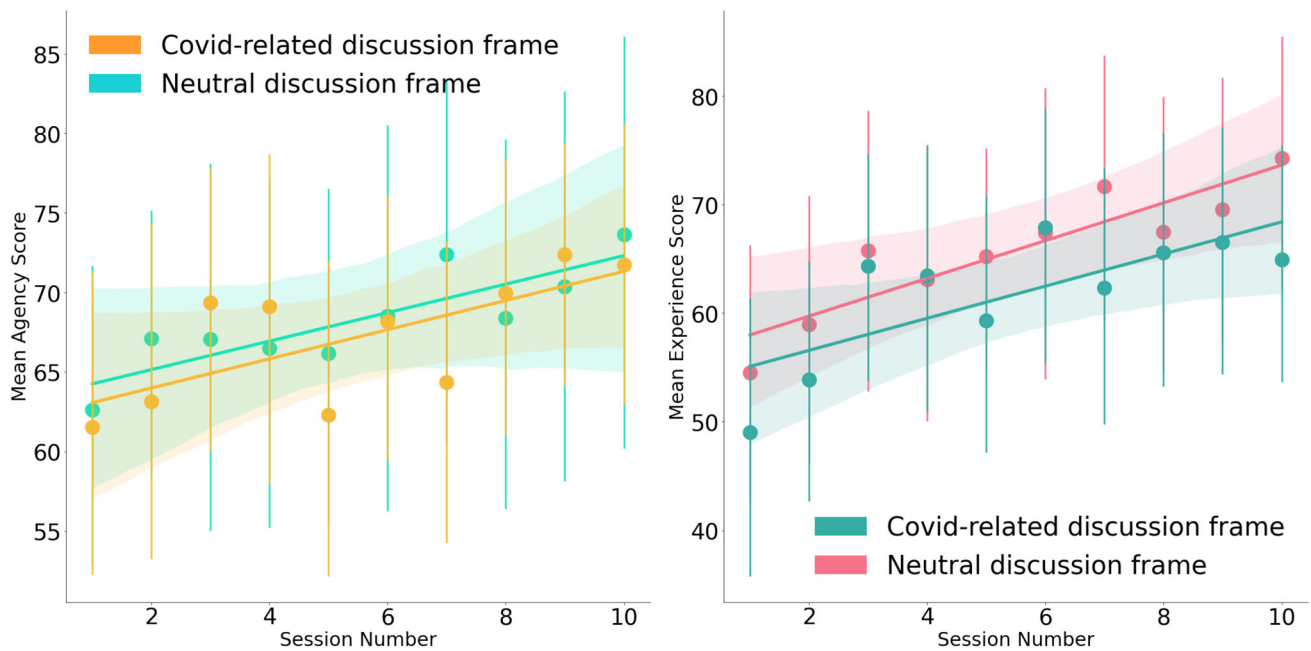
Fig. 6 From left to right: (1) Mean agency scores by session number and discussion frame. (2) Mean experience scores by session number and discussion frame

4.3.2 Comforting Responses

The model explains 78.3% of the variance in participants' perceptions of Pepper's *comforting responses*, whereas the fixed effects in the model explain 3.9% of the variance (see Table 5). The results stress that despite the variance between the participants (SD =.79), the session number has a significant positive fixed effect on participants' perceptions of Pepper's comforting responses (β =.07, SE =.01, $p < .001$, see Fig. 8). Nevertheless, no significant fixed effects emerged in terms of the discussion frame (β =.13, SE =.27, p =.635), and the interaction term of the session number and discussion frame ($\beta = -.01$, SE =.02, p =.540).

4.3.3 Loneliness

The model explains 75.9% of the variance in participants' feelings of *loneliness*, whereas the fixed effects in the model explain 7.8% of the variance (see Table 5). The results stress that despite the variance between the participants ($SD = 1.08$), the session number has a significant negative fixed effect on participants' feelings of loneliness ($\beta = -.05$, SE =.01, $p < .001$, see Fig. 9). Nevertheless, no significant fixed effects emerged in terms of the discussion frame (β =.63, SE =.37, p =.091), and the interaction term of the session number and discussion frame (β =.01, SE =.02, p =.674).

Another linear mixed effects model was used, omitting the data units collected in the induction session (session 0) before the exposure to the discussion frame manipulation, in
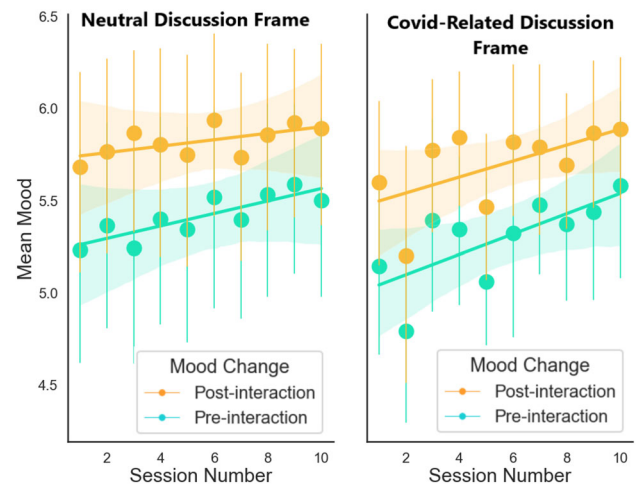

Fig. 7 From left to right: (1) Mean mood scores of participants in the neutral discussion frame, before and after the interaction, by session number. (2) Mean mood scores of participants in the Covid-related discussion frame, before and after the interaction, by session number
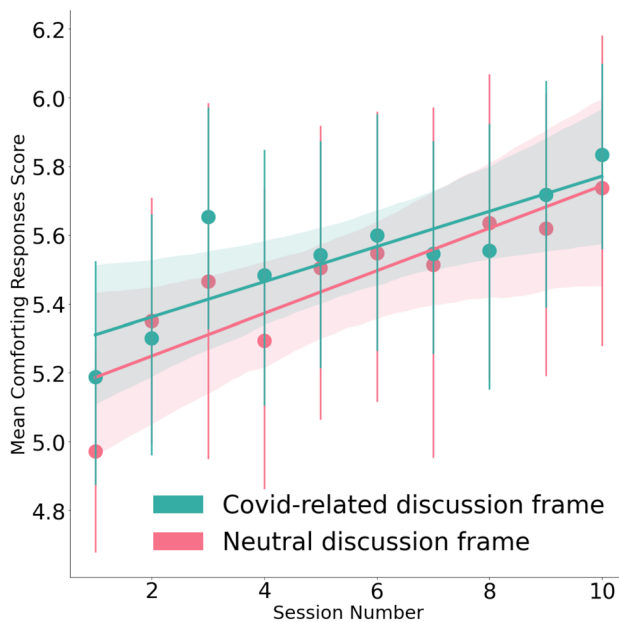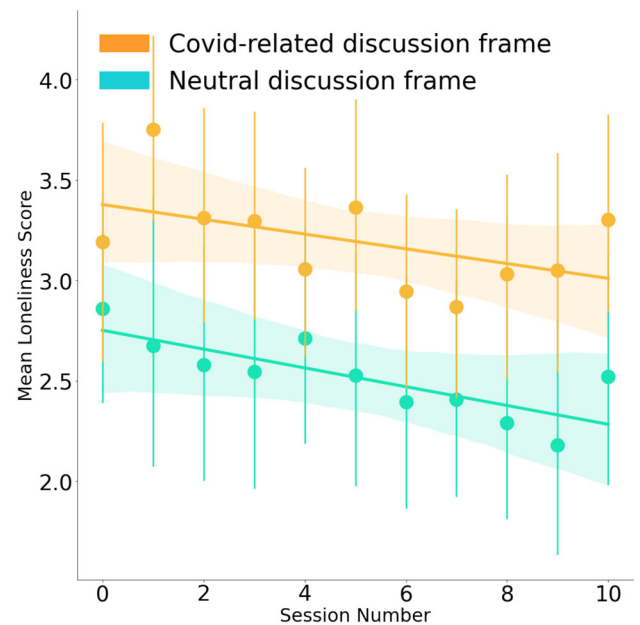
order to have a better evaluation of the effect of the discussion frame on participants' feelings of *loneliness*. The model explains 79.1% of the variance in participants' feelings of loneliness, whereas the fixed effects in the model explain 8.4% of the variance (see Table 5). The results stress that despite the variance between the participants ($SD = 1.10$), the session number has a significant negative fixed effect on participants' feelings of loneliness ($\beta = -.04$, SE =.02, $p = .008$), and the discussion frame has a significant fixed effect on participants' feelings of loneliness (β =.77, SE

Table 5 Results of well being

Fixed Effects	Mood		Comforting Responses		Loneliness		Loneliness (without session 0)		Stress	
	Estimates	*95%CI*	*Estimates*	*95%CI*	*Estimates*	*95%CI*	*Estimates*	*95%CI*	*Estimates*	*95%CI*
Intercept	5.23***	4.77 – 5.69	5.10***	4.72 – 5.47	2.76***	2.24 – 3.27	2.71***	2.18 – 3.24	2.70***	2.05 – 3.34
Discussion frame	-0.24	-0.87 – 0.40	.13	-.40 – .66	.63	-0.09 – 1.35	0.77*	0.03 – 1.51	1.31**	0.41 – 2.22
Session number	0.03*	0.01 – 0.06	0.07***	0.04 – 0.09	-0.05***	-0.08 – -0.02	-0.04**	-0.07 – -0.01	0.04	-0.02 – 0.11
Mood change	0.48***	0.27 – 0.70								
Discussion frame * Session number	0.02	-0.01 – 0.05	-0.01	-0.04 – 0.02	0.01	-0.03 – 0.05	-0.01	-0.05 – 0.03	-0.10*	-0.19 – -0.01
Discussion frame * Mood change	-0.01	-0.19 – 0.17								
Session number * Mood change	-0.01	-0.05 – 0.02								
Random Effects										
SD	0.94		0.79		1.08		1.10		0.86	
σ^2	0.41		0.18		0.41		0.36		0.25	
τ_{00}	0.89		0.62		1.16		1.21		0.74	
ICC	0.68		0.77		0.74		0.77		0.75	
N	39		39		39		39		39	
Observations	772		386		425		386		78	
Marginal R^2 / Conditional R^2	0.042 / 0.698		0.039 / 0.783		0.078 / 0.759		0.084 / 0.791		0.094 / 0.773	

Note: * $p<0.05$ ** $p<0.01$ *** $p<0.001$

For stress, 'Session number' is a dummy variable of sessions 5 and 10, with session 5 as the reference group (5 = 0).

Fig. 8 Mean comforting responses scores by session number and discussion frame

Fig. 9 Mean loneliness scores by session number and discussion frame

=.38, p = .046, see Fig. 9). Participants in the COVID-related experiences discussion frame group reported higher levels of loneliness compared to participants in the general experiences discussion frame group. Nevertheless, no significant fixed effect emerged in terms of the interaction term of the session number and discussion frame ($\beta = -.01$, SE =.02, p =.575) on participants' feelings of loneliness.

4.3.4 Stress

The model explains 77.3% of the variance in participants' feelings of *stress*, whereas the fixed effects in the model explain 9.4% of the variance (see Table 5). The results stress that despite the variance between the participants (SD =.86),

the discussion frame has a significant fixed effect on participants' feelings of stress ($\beta = 1.31$, SE =.45, p = .005). Participants in the COVID-related experiences discussion frame group reported higher levels of stress compared to participants in the general experiences discussion frame group. Moreover, the interaction term of the session number and discussion frame also has a significant fixed effect on participants' feelings of stress ($\beta = -.01$, SE =.05, p =.042) with participants in the COVID-related experiences discussion frame group reporting that their feelings of stress decreased from the fifth session to the tenth, whereas participants in the general experiences discussion frame group reported for increasing levels of stress from the fifth session to the tenth. Finally, no significant fixed effect emerged in terms of the

session number across the entire sample (β =.04, SE =.03, p =.205).

5 Discussion

Here we have introduced a novel long-term mediated experimental design to evaluate the extent to which a social robot can elicits and influence peoples' self disclosures to the robot, and how perceptions of the robot develop over time. Moreover, we measured the extent to which interactions with the social robot affected participants' well-being in different ways across time. Participants conversed with the social robot Pepper during 10 sessions distributed over 5 weeks, about one of two different topics depending on random group assignment. One group's conversation topics were framed within the context of the Covid-19 pandemic (e.g., social relationships during the pandemic, sustaining mental health during the pandemic, etc.), whereas the other group's conversation topics were similar, except no explicit mention of the Covid-19 pandemic was ever made. We evaluated the effect of time (session number) as well as how the discussion frame affected participants, comparing general everyday topics, to the same topics framed to the Covid-19 pandemic to address a more emotional context.

5.1 People Self-disclose Increasingly More to a Social Robot Over Time

Our first key finding shows that across the 10 sessions, people speak longer and share more information in their disclosures to the social robot Pepper. Moreover, consistent with previous results [3], subjective perceptions of self-disclosure align well with the objective data, and correspond to observed evidence of the length and duration of the disclosure, as people correctly perceived themselves to gradually share more information with Pepper across sessions. Finally, we found that people were more positive in their disclosures over time. The effects described here were even more meaningful when addressing only disclosures that are related to the session's conversation topic. Nevertheless, our results also reveal that the discussion frame has no meaningful nor significant effect on participants' disclosures to Pepper. Self-disclosure is a dynamic and socially complex human behaviour [12, 13], and accordingly, this key finding contributes to our understanding of humans' social behaviour and communication with robots. While numerous prior studies have exported humans' social behaviour towards robots in single-session studies, our knowledge of how people's behaviours towards robots change or develop over the longer-term remain limited in social HRI. Naturally, we recognize that people are different and might adapt different behavioural patterns when conversing with social robots. Nonetheless, we showed that

people self-disclosed increasingly more to Pepper over time in a systematic fashion even when the potential for such inter-individual differences are taken into account through the use of rigorous methodology. This is a meaningful contribution to HRI theory, showing that prolonged and intensive interactions with social robots can overcome novelty effects from behavioural objective evidence and not only from users' self-reported subjective perceptions.

5.2 People Perceive a Robot as More Social and Competent Over Time

We found that across the 10 sessions, participants attributed higher qualities of mind [see 101], in terms of agency and experience. Likewise, over time participants found Pepper to be friendlier and warm, as well as Pepper's communication skills more competent. Finally, across time, participants also rated the interactions with Pepper to be of increasingly higher quality. Here again, our results stress that the discussion frame has no meaningful nor significant effect on the way people perceive Pepper and the interaction. This key finding highlights the extent of people's social perception of robots over time. Despite Pepper's limited responses, over time participants attributed more social qualities to this particular robot, thus providing evidence for the influence of social engagement with a robot on its social perception over time. Furthermore, beyond finding Pepper to be more social, participants also attributed higher degrees of competency to Pepper over time. It is of note that people's perceptions of the robot and the interactions corresponded to their self-disclosure behaviour toward the robot over time. This key finding supports previous research showing how people's behaviours aligned with their social perceptions and attitudes towards the robot in single-session interactions [113]. Here we provided further support for this behavioural mechanism in HRI, and our results demonstrate how perceptions of robots and behaviours toward robots co-align over time during prolonged interactions.

5.3 Establishing Relationships with Social Robots

While previous longitudinal studies often report novelty effects in human–machine communication encounters [e.g., 52, 53], here we see a clear opposing trend, with evidence rooted in people's objective behaviour to robots (i.e., with the length and quality of participant disclosures increasing over time) and their subjective perceptions of robots (i.e., with participants' social perceptions of Pepper increasing over time, in terms of Pepper's agency, experience, friendliness and warmth, communication competency, and the interaction quality). These findings are particularly interesting as they provide clear evidence for social robots' potential to establish meaningful relationships with human users. While

consistent with previous suggestions on the matter [see 114, 115], the present study provides initial support for long-term relationships between humans users and a social robot, supported with multidimensional data. Furthermore, our findings establish important foundations for future HRI studies looking into how human-robot relationships develop over time, as well as for roboticists trying to create meaningful relationships between their robots and their users. Finally, these results highlight how human–robot relationships could act as ideal settings for robotic interventions for well-being. In addition, compared to Croes's previous studies [e.g., 52, 53] and despite our previous results in single-session studies [see 3], the present study suggests differences between embodied and disembodied agents in long-term interactions. We assume that people might attribute more social qualities to embodied agents (for the scope of this study, social robots) and accordingly, the relationship with such agents should evolve over time and not experience the same degree of novelty effects as experienced in [52, 53]. Nevertheless, this calls for further investigation and clear opportunities exist for future research to address the effects of embodiment on relationship establishment with artificial agents.

Our results further support the notion that the social dynamic in HRI, where humans often seek to establish social connections and rapport with robots, influences people's perceptions and attitudes towards the robot. The results of the present study further confirmed that mere-exposure effect [22] in HRI operates differently compared to traditional HCI, as the focus shifts from usability to the establishment of social bonds. By simulating human-like behaviours and engaging in social interactions, social robots, like the Pepper robot used in the present study, can elicit positive responses and be perceived as increasingly socially competent over time. This highlights the importance of understanding the dynamics of human–robot communication in long-term interactions and the potential for social robots to establish meaningful relationships with human users [8]. This distinction becomes evident when examining prolonged social interactions with a robot, resulting from repeated exposure. In the current context, we observe that these interactions demonstrate increasingly social behaviour and perception, representing the richest form of adaptation toward a social robot. Our findings suggest that users are not solely treating the Pepper robot used here as an object, but are willing to engage in long-term social interactions, and perhaps even establish some form of social connection with Pepper. Thus, our study highlights the difference between learning how to use an object through repeated usage, as observed in traditional HCI studies [see 23], and the social behaviour and perception exhibited toward a robot in HRI settings. This distinction emphasises the unique nature of social interaction in HRI and the need for a deeper understanding of human–robot communication beyond traditional usability perspectives.

5.4 Talking to Robots Positively Affects People's Well-Being

In terms of well-being, we found that participants' moods improved after interacting with Pepper, and also across the 10 sessions. Moreover, across the 10 sessions, participants reprted Pepper's responses to be more comforting. Our results revealed that the discussion frame per se did not have a meaningful or significant effect on people's moods and on the way people perceived Pepper's comforting responses. These findings provide further valuable evidence for the positive outcomes of employing a social robot as an intervention supporting people's well-being. Moreover, our results here add to previous studies [e.g., 38, 47, 116] that show the benefits of using robots for emotional support. Taken together with other results from this study (i.e., that people self-disclose increasingly more to a social robot over time and that people perceive a robot as more social over time), this study provides crucial evidence for establishing relationships with robots in health and care settings. These findings contribute to the introduction of social robots as conversational partners, and how this type of verbal interaction could support people with emotional regulation by talking about stressors and well-being. Simple tasks, like the one described in the study, are relatively easy to administer automatically in HRIs (by focusing on providing general and broad responses to users' disclosures) but can simulate effective procedures via self-disclosure like affect labelling [78] and other emotional introspective processes with users self-reflecting on their emotions and behaviours [79]. Accordingly, social robots can offer meaningful opportunities for self-managed interventions designed to support people's emotional health and well-being.

Another key finding in this regard has to do with people's feelings of loneliness. We found that over time across the experiment, participants reported feeling significantly less lonely. Loneliness is both a risk factor and a symptom of mental disorders and is a significant and growing public health issue with many comorbidities [117]. The recent COVID-19 pandemic stressed loneliness's tremendous effect on individuals' lives and society and highlighted the need for accessible intervention and support [118]. Social robots are often discussed as potential companions for people suffering from loneliness [see 119–121], especially concerning the Covid-19 pandemic [e.g., 63] with growing media attention [e.g., 122] and public initiatives [see 123]. Our results here further support that using objective and systematic measures, showing that repeated interactions with social robots reduced people's feelings of loneliness. This calls for further innovation and future research targeting loneliness as a public health issue using social robots.

5.5 Robots that Discuss Emotional Content can Simulate Feelings

Consistent with previous results in single-session HRIs [3], the discussion frame did not affect people's self-disclosure toward social robots or the way they perceived the robot or the interaction. However, our results do suggest that framing a discussion with a robot around a more emotional topic may elicit more emotional feelings among participants. This was specifically observed in this study with feelings of loneliness and stress. Our results here showed that when Pepper addressed the COVID-19 pandemic, participants reported higher levels of loneliness and stress, compared to participants in the general experiences discussion frame group. This important finding provides initial support for the notion that robots can trigger an emotional reaction from the interaction's content. When studying robots' affective capabilities, previous studies often address factors related to the robot's visual features (e.g., embodiment) or robotic functionalities (e.g., emotional recognition) [124]. Yet, studies aiming at developing and assessing social robotic interventions for well-being should also study the robot's ability to simulate human affect in different ways [see 124]. Our results highlight the role of content and frame when aiming to simulate human emotions and feelings during HRIs. they further show that robots can trigger complex emotions when addressing meaningful and personal moments and events. Nonetheless, our evidence here is based solely on two factors of loneliness and stress, answering to one emotional frame—mentioning the Covid-19 pandemic. Thus, for further understanding humans' emotional response to social robotic stimuli, this should be studied with various feelings and emotions, within several settings and in response to different frames. It is important to acknowledge that the limited differences observed between the conditions in this study may be attributed to ceiling effects. This could be because the study took place during the peak of the pandemic, causing participants to primarily focus on the consequences of the pandemic regardless of their assigned condition. Therefore, future studies could explore other approaches for manipulating emotional themes during experimental social interactions with social robots. While established emotion-elicitation techniques have long been used in human–human social interaction research [125], it may be challenging to seamlessly integrate them into HRI behavioural paradigms. Therefore, we believe that researchers interested in such questions further explore these techniques and evaluate their capacity to evoke varying levels of emotion in social interactions with robots.

5.6 Methodological Contribution

Through the present research, we aimed to establish experimental methods that researchers from HRI, as well as from a number of related fields, including psychology, psychiatry, social work, anthropology, and computer science, might wish to use to further explore people's perceptions of a sociable, humanoid robot in natural everyday settings during prolonged conversational interactions. Beyond exploring general questions regarding how people engage with a social robot from their home settings and how it supports their well-being, the current research also provides a means to further examine the impact of novelty effects, and the impact of long-term social engagement with a robot on behaviour [c.f., 56]. Furthermore, this study can be replicated and tested with various populations, clinical and healthy, in order to understand how social robots could be introduced in different care settings and as interventions using speech-based interactions [c.f., 126]. By introducing this novel paradigm in detail here, and documenting results from a rigorous empirical study using this paradigm, we aim to provide a tool that we hope will be of use to the HRI research community more broadly, while also assisting with facilitating research rigour and reproducibility [1, 8, 127, 128], as well as the development of data-centric robotic models [c.f., 129–131]. Moreover, we would argue that the online computer-mediated means of human-robot communication used in this experimental design can overcome some of the challenges and barriers that are related to long-term HRI studies in natural ecologically valid settings (such as the costs associated with sending individual robots home for an extended period of time with participants) and suggest alternative means for conducting HRI research in people's natural settings.

6 Limitations and Future Research

Our study has contributed valuable insights into the effects of long-term repeated interactions with a social robot on self-disclosure behaviour, perceptions, and well-being. However, several limitations should be acknowledged, which open up avenues for future research to deepen our understanding of HRIs and their potential applications in supporting emotional well-being.

6.1 Mediated Embodiment Limits Users' Perception of Robots

Due to the mediated nature of the interactions, participants' perception of Pepper's embodiment and physical presence may have been limited. Conducting the study online enabled us to reach a larger and more diverse sample size, enhancing the external validity of our findings while being cost-effective

[132]. This method has proven valuable in generating insights and hypotheses that can later be further examined in real-life settings, enabling a more comprehensive understanding of the topic. Furthermore, while some previous studies claim for the moderating role of physical embodiment [133, 134], recent experimental studies that compared in-person interactions with mediated interactions involving social robots have reported no significant differences in participants' perception and behaviour [3, 135, 136]. Although online settings may not fully replicate real-life interactions with social robots, they provide an initial exploration of the potential effects of long-term interactions and allow us to examine the specific research questions we aimed to investigate. This is particularly significant due to the widespread adoption of CMC during the Covid-19 pandemic, which made online interactions more commonplace and therefore made our experiment more reflective of the prevailing social context [68]. The controlled environment of an online experiment also facilitated consistent conditions across participants and minimised confounding variables, which is essential for drawing reliable conclusions. While our study's outcomes offer significant benchmarks and valuable insights for future investigations conducted in real-life situations, they can also provide insights into the significance of robots' physical presence compared to the prolonged mediated interactions observed in our study. To address the limitation of generalizability, future research could incorporate real-life interactions with social robots to validate and extend our findings. By comparing outcomes from online and in-person interactions, researchers can gain insights into how embodiment influences the effectiveness of social robots as conversational partners.

6.2 The Absence of a Control group

Another limitation of our study is the absence of a control group. Including a control group would have strengthened our ability to make direct comparisons and isolate the effects of the social robot [137]. However, we employed a mixed-factorial design with repeated interactions (i.e., measures), where each participant served as their own control, allowing us to examine changes within individuals over time. This design provided a baseline for comparison and enabled us to examine changes within individuals over time through repeated interactions with the social robot [138, 139]. Moreover, the logistical challenges associated with recruiting and managing an additional control group were considered, which influenced our decision for this experimental design. Whilst the absence of a control group limits the strength of our causal claims to the presence of the robot, our repeated measures design provides valuable insights into the changes occurring over time in individuals' self-disclosure behaviour, perceptions, and well-being during long-term interactions with a social robot. Future research incorporating a control

group and investigating additional variables or interventions alongside the social robot could provide a more rigorous comparison and further disentangle the specific impacts of the robot itself.

6.3 The Challenges of Measuring Self Disclosure and Well-Being

The study primarily relied on self-reported and behavioural measures rather than physiological indicators to assess participants' well-being. While recognizing the value of incorporating physiological measures, logistical constraints posed by conducting the experiment via Zoom and the long-term nature of the study required careful prioritisation of research objectives. Self-report measures were chosen as they are widely accepted for assessing subjective experiences and have been extensively used in well-being research [140–142]. Despite limitations, the focus on self-report measures offers valuable insights into participants' subjective well-being experiences during long-term interactions with a social robot. Moreover, previous studies found that self reported measurements aligned with participants objective behaviour [140–142], also in HRI research [3, 113] and in regards to self-disclosure behaviour [3]. Future studies can replicate the experimental design while incorporating additional physiological measures to further assess participants' well-being and emotional changes over time. This approach would enable researchers to build upon the findings from this study, and explore the interplay between affective interactions with social robots, well-being, and objective physiological and behavioural indicators.

Further analysis of the disclosure content through automated and manual text analysis methods could provide deeper insights into the emotional content conveyed in participants' disclosures to the robot. Conducting a more comprehensive textual analysis using advanced computational techniques [see 143, 144] and manual evaluation [e.g., 145–147] could offer valuable insights and context to participants' self-disclosures to the robot [93]. Although the study did not utilise manual coding of self-disclosure, the behavioural paradigm employed was specifically designed to encourage self-disclosure by asking participants questions about personal matters. The analysis of both all items and items related to the topic of disclosure provides a comprehensive view of self-disclosure behaviour observed in the study. This approach allows for exploration of broader patterns and trends in self-disclosure and an examination of specific instances of personal disclosures within the defined topics. While manually coding each instance of self-disclosure may be challenging due to the substantial amount of data, using basic measurements such as duration and word count enhances reliability and consistency. These measurements objectively assess changes in participants' behaviour

over time and correlate with personal and meaningful self-disclosures as observed in previous research [95–97].

In future research, the open-ended questions incorporated in the study should be analysed using qualitative methods to deepen the understanding of human disclosures to social robots and explore subjective responses provided by participants [89]. Additionally, qualitatively analysing the content of the disclosures will allow for a more in-depth exploration of the nuances and frames within the self-disclosures. Furthermore, conducting secondary analyses to examine the effect of disclosure topics could provide valuable insights [c.f., 3]. This exploration could shed light on the effects of different topics and the order of presentation, contributing valuable knowledge to the field.

7 Conclusion

These results set the stage for addressing social robots as conversational partners in social settings, and how this type of verbal interaction could support people with emotional regulation by talking about stressors and well-being. The study provides crucial evidence for establishing relationships with robots, and their potential introduction as interventions supporting people's emotional health through encouraging self-disclosure. These results provide meaningful evidence for user experience, acceptance, and trust of social robots and other conversational agents [148, 149], highlighting how the perception of robots and behaviour towards them is closely related. These results hold several implications for assessing interactions as well as interventions with socially assistive robots, and for HRI research in general. Future research is encouraged to replicate and reproduce the current findings with different robots and different populations. In doing so, this will help to overcome the vast challenges and barriers that are related to long-term HRI studied in natural ecologically valid settings.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Statement The study followed rigorous ethical standards and all study procedures were approved by the research ethics of the University of Glasgow (Ethics Approval Numbers 300200094 & 300200132). All participants provided written informed consent before participating in the study, and were debriefed after participating in the study (for more information see Sect. 3.2.4).

References

1. Henschel A, Laban G, Cross ES (2021) What makes a robot social? A Review of social robots from science fiction to a home or hospital near you. Curr Robot Rep 2:9–19. https://doi.org/10.1007/s43154-020-00035-0

2. Hortensius R, Hekele F, Cross ES (2018) The perception of emotion in artificial agents. IEEE Trans Cognit Dev Syst 10(4):852–864. https://doi.org/10.1109/TCDS.2018.2826921

3. Laban G, George J-N, Morrison V, Cross ES (2021) Tell me more! Assessing interactions with social robots from speech. Paladyn J Behav Robot 12(1):136–159. https://doi.org/10.1515/pjbr-2021-0011

4. Smedegaard CV (2019) Reframing the role of novelty within social HRI: from noise to information. In: Proceedings of the 14th ACM/IEEE international conference on human-robot interaction. HRI'19. IEEE Press, pp 411–420. https://doi.org/10.1109/HRI.2019.8673219

5. Smedegaard CV (2022) Novelty knows no boundaries: why a proper investigation of novelty effects within SHRI should begin by addressing the scientific plurality of the field. Front Robot AI. https://doi.org/10.3389/FROBT.2022.741478

6. Breazeal C (2003) Toward sociable robots. Robot Auton Syst 42(3):167–175. https://doi.org/10.1016/S0921-8890(02)00373-1

7. Robinson NL, Cottier TV, Kavanagh DJ (2019) Psychosocial health interventions by social robots: systematic review of randomized controlled trials. J Med Internet Res 21(5):1–20. https://doi.org/10.2196/13203

8. Cross ES, Ramsey R (2021) Mind meets machine: towards a cognitive science of human-machine interactions. Trends Cogn Sci 25(3):200–212. https://doi.org/10.1016/J.TICS.2020.11.009

9. Hortensius R, Cross ES (2018) From automata to animate beings: the scope and limits of attributing socialness to artificial agents. Ann N Y Acad Sci 1426(1):93–110. https://doi.org/10.1111/nyas.13727

10. Cross ES, Ramsey R, Liepelt R, Prinz W, Hamilton AFdC (2016) The shaping of social perception by stimulus and knowledge cues to human animacy. Philos Trans R Soc Lond Ser B Biol Sci 371(1686):20150075. https://doi.org/10.1098/rstb.2015.0075

11. Laban G (2022) Social robots as communication partners to support emotional health and well-being. In: 2022 10th International conference on affective computing and intelligent interaction workshops and demos (ACIIW), pp. 1–5. IEEE, Nara, Japan. https://doi.org/10.1109/ACIIW57231.2022.10086018. https://ieeexplore.ieee.org/document/10086018/

12. Jourard SM, Lasakow P (1958) Some factors in self-disclosure. Psychol Sci Public Interest 56(1):91–98. https://doi.org/10.1037/h0043357

13. Pearce WB, Sharp SM (1973) Self-disclosing communication. J Commun 23(4):409–425. https://doi.org/10.1111/j.1460-2466.1973.tb00958.x

14. Frattaroli J (2006) Experimental disclosure and its moderators: a meta-analysis. Psychol Bull 132(6):823–865. https://doi.org/10.1037/0033-2909.132.6.823

15. Frisina PG, Borod JC, Lepore SJ (2004) A meta-analysis of the effects of written emotional disclosure on the health outcomes of clinical populations. J Nervous Mental Disease 192(9)

16. Kennedy-Moore E, Watson JC (2001) How and when does emotional expression help? Rev Gen Psychol 5(3):187–212. https://doi.org/10.1037/1089-2680.5.3.187

17. Sloan DM (2010) Self-disclosure and psychological well-being. Social psychological foundations of clinical psychology, pp. 212–225. The Guilford Press, New York

18. Derlega VJ, Winstead BA, Lewis RJ, Maddux J (1993) Clients' responses to dissatisfaction in psychotherapy: a test of Rusbult's exit-voice-loyalty-neglect model. J Soc Clin Psychol 12(3):307–318. https://doi.org/10.1521/jscp.1993.12.3.307

19. Colquhoun HL, Squires JE, Kolehmainen N, Fraser C, Grimshaw JM (2017) Methods for designing interventions to change healthcare professionals' behaviour: a systematic review. Implement Sci 12(1):30. https://doi.org/10.1186/s13012-017-0560-5

20. Wight D, Wimbush E, Jepson R, Doi L (2016) Six steps in quality intervention development (6SQuID). J Epidemiol Community Health 70(5):520. https://doi.org/10.1136/jech-2015-205952

21. Kappas A, Stower R, Vanman EJ (2020) Communicating with robots what we do wrong and what we do right in artificial social intelligence, and what we need to do better. In: Sternberg RJ, Kostic A (eds) Communication Nonverbal Intelligence Social. Springer, Cham, pp 233–254. https://doi.org/10.1007/978-3-030-34964-6

22. Zajonc RB (2001) Mere exposure: a gateway to the subliminal. Curr Dir Psychol Sci 10(6):224–228. https://doi.org/10.1111/1467-8721.00154

23. Van Giesen RI, Fischer ARH, Van Dijk H, Van Trijp HCM (2015) Affect and cognition in attitude formation toward familiar and unfamiliar attitude objects. PLoS ONE. https://doi.org/10.1371/JOURNAL.PONE.0141790

24. Issa T, Isaias P (2015) Usability and Human Computer Interaction (HCI). Sustain Design. https://doi.org/10.1007/978-1-4471-6753-2

25. Polakow T, Laban G, Teodorescu A, Busemeyer JR, Gordon G (2022) Social robot advisors: effects of robot judgmental fallacies and context. Intel Serv Robot 15(5):593–609. https://doi.org/10.1007/s11370-022-00438-2

26. Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: a survey. Int J Soc Robot 5(2):291–308. https://doi.org/10.1007/s12369-013-0178-y

27. Sabanovic S, Michalowski MP, Simmons R (2006) Robots in the wild: observing human-robot social interaction outside the lab. Int Workshop Adv Motion Control AMC 2006:576–581. https://doi.org/10.1109/AMC.2006.1631758

28. Dautenhahn K (2007) Socially intelligent robots: dimensions of human-robot interaction. Philos Trans R Soc Lond Ser B Biol Sci 362(1480):679–704. https://doi.org/10.1098/rstb.2006.2004

29. Sung J, Christensen HI, Grinter RE (2009) Robots in the wild: understanding long-term use. In: 2009 4th ACM/IEEE international conference on human-robot interaction (HRI), pp. 45–52. ACM, New York

30. Stower R, Kappas A (2020) Oh no, my instructions were wrong!" An exploratory pilot towards children's trust in social robots. In: 2020 29th IEEE international conference on robot and human interactive communication (RO-MAN), pp. 641–646. https://doi.org/10.1109/RO-MAN47096.2020.9223495

31. Holthaus P, Wachsmuth S (2021) It was a pleasure meeting you: towards a holistic model of human-robot encounters. Int J Soc Robot. https://doi.org/10.1007/s12369-021-00759-9

32. Céspedes N, Irfan B, Senft E, Cifuentes CA, Gutierrez LF, Rincon-Roncancio M, Belpaeme T, Múnera M (2021) A socially assistive robot for long-term cardiac rehabilitation in the real world. Front Neurorobot 15:21. https://doi.org/10.3389/FNBOT.2021.633248

33. De Graaf MMA, Allouch SB, Klamer T (2015) Sharing a life with Harvey: exploring the acceptance of and relationship-building with a social robot. Comput Hum Behav 43:1–14. https://doi.org/10.1016/J.CHB.2014.10.030

34. De Graaf MMA, Allouch SB, van Dijk JAGM (2016) Long-term evaluation of a social robot in real homes. Interact Stud 17(3):462–491

35. Feingold Polak R, Tzedek SL (2020) Social robot for rehabilitation: expert clinicians and post-stroke patients' evaluation following a long-term intervention. In: Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction. HRI'20, pp. 151–160. Association for Computing Machinery, New York. https://doi.org/10.1145/3319502.3374797

36. Levinson L, Gvirsman O, Gorodesky IM, Perez A, Gonen E, Gordon G (2020) Learning in summer camp with social robots: a morphological study. Int J Soc Robot. https://doi.org/10.1007/s12369-020-00689-y

37. van Maris A, Zook N, Caleb-Solly P, Studley M, Winfield A, Dogramadzi S (2020) Designing ethical social robots-a longitudinal field study with older adults. Front Robot AI 7:1. https://doi.org/10.3389/FROBT.2020.00001

38. Jeong S, Aymerich-Franch L, Arias K, Alghowinem S, Lapedriza A, Picard R, Park HW, Breazeal C (2022) Deploying a robotic positive psychology coach to improve college students' psychological well-being. User Model User-Adap Inter. https://doi.org/10.1007/s11257-022-09337-8

39. Kanda T, Hirano T, Eaton D, Ishiguro H (2004) Interactive robots as social partners and peer tutors for children: a field trial. Human-Comput Interact 19(1):61–84

40. Michaelis JE, Mutlu B (2017) Someone to read with: design of and experiences with an in-home learning companion robot for reading. In: Proceedings of the 2017 CHI conference on human factors in computing systems. https://doi.org/10.1145/3025453

41. Michaelis JE, Mutlu B (2019) Supporting interest in science learning with a social robot. In: Proceedings of the 18th ACM international conference on interaction design and children. https://doi.org/10.1145/3311927

42. Woo H, LeTendre GK, Pham-Shouse T, Xiong Y (2021) The use of social robots in classrooms: a review of field-based studies.

Educ Res Rev 33:100388. https://doi.org/10.1016/j.edurev.2021.100388

43. Robinson NL, Ward B, Kavanagh DJ (2021) A robot-delivered program for low-intensity problem-solving therapy for students in higher education. In: 2021 30th IEEE international conference on robot and human interactive communication, RO-MAN 2021, pp. 945–950. https://doi.org/10.1109/RO-MAN50785.2021.9515532

44. Robinson NL, Connolly J, Hides L, Kavanagh DJ (2020) A social robot to deliver an 8-week intervention for diabetes management: initial test of feasibility in a hospital clinic. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 12483 LNAI, pp 628–639. https://doi.org/10.1007/978-3-030-62056-1_

45. Robinson NL, Connolly J, Hides L, Kavanagh DJ (2020) Social robots as treatment agents: pilot randomized controlled trial to deliver a behavior change intervention. Internet Interv 21:100320. https://doi.org/10.1016/j.invent.2020.100320

46. Nomura T, Kanda T, Yamada S, Suzuki T (2021) The effects of assistive walking robots for health care support on older persons: a preliminary field experiment in an elder care facility. Intell Serv Robot 14(1):25–32. https://doi.org/10.1007/S11370-020-00345-4

47. Bodala IP, Churamani N, Gunes H (2021) Teleoperated robot coaching for mindfulness training: a longitudinal study. In: 2021 30th IEEE international conference on robot and human interactive communication, RO-MAN 2021, pp 939–944. https://doi.org/10.1109/RO-MAN50785.2021.9515371

48. Céspedes N, Raigoso D, Múnera M, Cifuentes CA (2021) Long-term social human-robot interaction for neurorehabilitation: robots as a tool to support gait therapy in the pandemic. Front Neurorobot 15:10. https://doi.org/10.3389/FNBOT.2021.612034

49. Koren Y, Feingold Polak R, Levy-Tzedek S (2022) Extended interviews with stroke patients over a long-term rehabilitation using human-robot or human-computer interactions. Int J Soc Robot 14(8):1893–1911. https://doi.org/10.1007/S12369-022-00909-7/FIGURES/6

50. Feingold-Polak R, Barzel O, Levy-Tzedek S (2021) A robot goes to rehab: a novel gamified system for long-term stroke rehabilitation using a socially assistive robot-methodology and usability testing. J Neuroeng Rehabil 18(1):122. https://doi.org/10.1186/s12984-021-00915-2

51. Afyouni A, Ocnarescu I, Cossin I, Kamoun E, Mazel A, Fattal C (2022) Living one week with an autonomous Pepper in a rehabilitation center: lessons from the field. In: RO-MAN 2022-31st IEEE international conference on robot and human interactive communication: social, asocial, and antisocial robots, pp 554–559. https://doi.org/10.1109/RO-MAN53752.2022.9900640

52. Croes EAJ, Antheunis ML (2020) Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. J Soc Pers Relat 38(1):279–300. https://doi.org/10.1177/0265407520959463

53. Croes EAJ, Antheunis ML (2021) 36 Questions to loving a chatbot: are people willing to self-disclose to a chatbot? In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol. 12604 LNCS, pp 81–95. Springer. https://doi.org/10.1007/978-3-030-68288-0

54. Ho A, Hancock J, Miner AS (2018) Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. J Commun 68(4):712–733. https://doi.org/10.1093/joc/jqy026

55. Hortensius R, Chaudhury B, Hoffmann M, Cross E (2022) Tracking human interactions with a commercially-available robot over multiple days. Open Res Europe 2:97. https://doi.org/10.12688/OPENRESEUROPE.14824.1

56. Cross ES, Riddoch KA, Pratts J, Titone S, Chaudhury B, Hortensius R (2019) A neurocognitive investigation of the impact of socializing with a robot on empathy for pain. Philos Trans R Soc B Biol Sci 374(1771):20180034. https://doi.org/10.1098/rstb.2018.0034

57. Scoglio AAJ, Reilly ED, Gorman JA, Drebing CE (2019) Use of social robots in mental health and well-being research: systematic review. J Med Internet Res 21(7):13322. https://doi.org/10.2196/13322

58. Laban G, Ben-Zion Z, Cross ES (2022) Social robots for supporting post-traumatic stress disorder diagnosis and treatment. Front Psychiatry. https://doi.org/10.3389/FPSYT.2021.752874

59. Nomura T, Kanda T, Suzuki T, Yamada S (2020) Do people with social anxiety feel anxious about interacting with a robot? AI Soc 35(2):381–390. https://doi.org/10.1007/s00146-019-00889-9

60. Rasouli S, Gupta G, Nilsen E, Dautenhahn K (2022) Potential applications of social robots in robot-assisted interventions for social anxiety. Int J Soc Robot 14(5):1–32. https://doi.org/10.1007/S12369-021-00851-0

61. Axelsson M, Churamani N, Caldir A, Gunes H (2022) Participant perceptions of a robotic coach conducting positive psychology exercises: a systematic analysis. https://doi.org/10.48550/arxiv.2209.03827

62. Robinson NL, Kavanagh DJ (2021) A social robot to deliver a psychotherapeutic treatment: qualitative responses by participants in a randomized controlled trial and future design recommendations. Int J Hum Comput Stud 155:102700. https://doi.org/10.1016/J.IJHCS.2021.102700

63. Scassellati B, Vázquez M (2020) The potential of socially assistive robots during infectious disease outbreaks. Sci Robot. https://doi.org/10.1126/scirobotics.abc9014

64. Henschel A, Cross ES (2020) The neuroscience of loneliness—and how technology is helping us. Convers 3–7

65. Yang G-Z, Nelson BJ, Murphy RR, Choset H, Christensen H, Collins SH, Dario P, Goldberg K, Ikuta K, Jacobstein N, Kragic D, Taylor RH, McNutt M (2020) Combating COVID-19-The role of robotics in managing public health and infectious diseases. Sci Robot 5(40):5589. https://doi.org/10.1126/SCIROBOTICS.ABB5589

66. Gasteiger N, Loveys K, Law M, Broadbent E (2021) Friends from the future: a scoping review of research into robots and computer agents to combat loneliness in older people. Clin Interv Aging 16:941–971. https://doi.org/10.2147/CIA.S282709

67. Isabet B, Pino M, Lewis M, Benveniste S, Rigaud A-S (2021) Social telepresence robots: a narrative review of experiments involving older adults before and during the COVID-19 pandemic. Int J Environ Res Public Health. https://doi.org/10.3390/IJERPH18073597

68. Choi M, Choung H (2021) Mediated communication matters during the COVID-19 pandemic: the use of interpersonal and masspersonal media and psychological well-being. J Soc Pers Relat 38(8):2397–2418. https://doi.org/10.1177/02654075211029378

69. Shiomi M, Nakata A, Kanbara M, Hagita N (2020) Robot reciprocation of hugs increases both interacting times and self-disclosures. Int J Soc Robot. https://doi.org/10.1007/s12369-020-00644-x

70. Duan Y, Yoon M, Liang Z, Hoorn JF (2021) Self-disclosure to a robot: only for those who suffer the most. Robotics 10(3):98. https://doi.org/10.3390/ROBOTICS10030098

71. Hoffman G, Birnbaum GE, Vanunu K, Sass O, Reis HT (2014) Robot responsiveness to human disclosure affects social impression and appeal. In: Proceedings of the 2014 ACM/IEEE international conference on human-robot interaction. HRI-14. Association for Computing Machinery, New York, pp 1–8. https://doi.org/10.1145/2559636.2559660

72. Nakamura Y, Umemuro H (2022) Effect of robot's listening attitude change on self-disclosure of the elderly. Int J Soc Robot. https://doi.org/10.1007/s12369-022-00934-6

73. De Groot J-J, Barakova E, Lourens T, van Wingerden E, Sterkenburg P (2019) Game-based human-robot interaction promotes self-disclosure in people with visual impairments and intellectual disabilities BT—understanding the brain function and emotions. In: Vicente JM, Álvarez-Sánchez JR, de la Paz López F, Toledo Moreo J, Adeli H (eds) Ferrández. Springer, Cham, pp 262–272

74. van Wingerden E, Barakova E, Lourens T, Sterkenburg PS (2020) Robot-mediated therapy to reduce worrying in persons with visual and intellectual disabilities. J Appl Res Intell Disabil. https://doi.org/10.1111/jar.12801

75. Pennebaker JW (1997) Writing about emotional experiences as a therapeutic process. Psychol Sci 8(3):162–166

76. Pennebaker JW, Beall SK (1986) Confronting a traumatic event: toward an understanding of inhibition and disease. J Abnorm Psychol 95(3):274–281. https://doi.org/10.1037//0021-843X.95.3.274

77. Luo RL, Zhang TXY, Chen DH-C, Hoorn JF, Huang IS (2022) Social robots outdo the not-so-social media for self-disclosure: safe machines preferred to unsafe humans? Robotics 11(5):92. https://doi.org/10.3390/ROBOTICS11050092

78. Torre JB, Lieberman MD (2018) Putting feelings into words: affect labeling as implicit emotion regulation. Emot Rev 10(2):116–124. https://doi.org/10.1177/1754073917742706

79. Tamir DI, Mitchell JP (2012) Disclosing information about the self is intrinsically rewarding. Proc Natl Acad Sci USA 109(21):8038–8043. https://doi.org/10.1073/pnas.1202129109

80. Creswell JD, Way BM, Eisenberger NI, Lieberman MD (2007) Neural correlates of dispositional mindfulness during affect labeling. Psychosom Med 690:560–565. https://doi.org/10.1097/PSY.0b013e3180f6171f

81. Laban G, Cross ES (2023) Sharing with robots: why do we do it and how does it make us feel? PsyArxiv. https://doi.org/10.31234/OSF.IO/2AZPQ

82. Nelson LD, Simmons JP, Simonsohn U (2012) Let's publish fewer papers. Psychol Inq 23(3):291–293. https://doi.org/10.1080/1047840X.2012.705245

83. Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychol Sci 22(11):1359–1366. https://doi.org/10.1177/0956797611417632

84. Laban G, Kappas A, Morrison V, Cross ES (2021) Socially assistive robots in prolonged interactions for measuring expression in natural settings. Technical report. https://doi.org/10.17605/OSF.IO/R5XGA. https://osf.io/r5xga

85. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, du Sert NP, Simonsohn U, Wagenmakers E-J, Ware JJ, Ioannidis JPA (2017) A manifesto for reproducible science. Nat Hum Behav 1(1):21. https://doi.org/10.1038/s41562-016-0021

86. Laban G, Kappas A, Morrison V, Cross ES (2020) OSF repository: socially assistive robots in prolonged interactions for measuring expression in natural settings. OSF. https://doi.org/10.17605/OSF.IO/M74CB. https://osf.io/m74cb/

87. Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 39(2):175–191. https://doi.org/10.3758/BF03193146

88. Faul F, Erdfelder E, Buchner A, Lang A-G (2009) Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 41(4):1149–1160. https://doi.org/10.3758/BRM.41.4.1149

89. Riddoch KA, Cross ES (2021) "Hit the robot on the head with this mallet" - making a case for including more open questions in HRI research. Front Robot AI 8:2. https://doi.org/10.3389/frobt.2021.603510

90. Aron A, Melinat E, Aron EN, Vallone RD, Bator RJ (1997) The experimental generation of interpersonal closeness: a procedure and some preliminary findings. Pers Soc Psychol Bull 23(4):363–377. https://doi.org/10.1177/0146167297234003

91. Jourard SM (1971) Self-disclosure: an experimental analysis of the transparent self. Wiley, Oxford, p 248248

92. Connell J, Brazier J, O'Cathain A, Lloyd-Jones M, Paisley S (2012) Quality of life of people with mental health problems: a synthesis of qualitative research. Health Qual Life Outcomes 10(1):138. https://doi.org/10.1186/1477-7525-10-138

93. Kreiner H, Levi-Belz Y (2019) Self-disclosure here and now: combining retrospective perceived assessment with dynamic behavioral measures. Front Psychol 10:558. https://doi.org/10.3389/fpsyg.2019.00558

94. Levi-Belz Y, Kreiner H (2016) What you say and how you say it: analysis of speech content and speech fluency as predictors of judged self-disclosure. Soc Psychol Personal Sci 7(3):232–239. https://doi.org/10.1177/1948550616632575

95. Barak A, Gluck-Ofri O (2007) Degree and reciprocity of self-disclosure in online forums. Cyber Psychol Behav 10(3):407–417. https://doi.org/10.1089/CPB.2006.9938

96. Joinson AN (2001) Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity. Eur J Soc Psychol 31(2):177–192. https://doi.org/10.1002/EJSP.36

97. Pedersen DM, Breglio VJ (1968) Personality correlates of actual self-disclosure. Psychol Rep 22(2):495–501. https://doi.org/10.2466/PR0.1968.22.2.495

98. Jadoul Y, Thompson B, de Boer B (2018) Introducing parselmouth: a python interface to praat. J Phon 71:1–15. https://doi.org/10.1016/j.wocn.2018.07.001

99. Boersma P, Weenink D (2001) PRAAT, a system for doing phonetics by computer. Glot Int 5:341–345

100. Hutto CJ, Gilbert E (2014) VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the 8th international AIII conference on weblogs and social media, ICWSM 2014, pp 216–225. AAAI, Ann Arbor, MI. https://doi.org/10.1609/icwsm.v8i1.14550

101. Gray HM, Gray K, Wegner DM (2007) Dimensions of mind perception. Science 315(5812):619–619. https://doi.org/10.1126/science.1134475

102. Petty RE, Mirels HL (1981) Intimacy and scarcity of self-disclosure: effects on interpersonal attraction for males and females. Pers Soc Psychol Bull 7(3):493–503. https://doi.org/10.1177/014616728173020

103. Birnbaum GE, Mizrahi M, Hoffman G, Reis HT, Finkel EJ, Sass O (2016) What robots can teach us about intimacy: the reassuring effects of robot responsiveness to human disclosure. Comput Hum Behav 63:416–423. https://doi.org/10.1016/j.chb.2016.05.064

104. Demeure V, Niewiadomski R, Pelachaud C (2011) How is believability of a virtual agent related to warmth, competence, personification, and embodiment? presence: teleoperators and virtual environments. Presence 20(5):431–448. https://doi.org/10.1162/PRES_a_

105. Berry DS, Hansen JS (2000) Personality, nonverbal behavior, and interaction quality in female dyads. Pers Soc Psychol Bull 26(3):278–292. https://doi.org/10.1177/0146167200265002

106. Nahum M, Van Vleet TM, Sohal VS, Mirzabekov JJ, Rao VR, Wallace DL, Lee MB, Dawes H, Stark-Inbar A, Jordan JT, Biagianti B, Merzenich M, Chang EF (2017) Immediate mood scaler: tracking symptoms of depression and anxiety using a novel mobile mood scale. JMIR Mhealth Uhealth 5(4):44. https://doi.org/10.2196/mhealth.6544

107. Crawford JR, Henry JD (2004) The positive and negative affect schedule (PANAS): construct validity, measurement properties

and normative data in a large non-clinical sample. Br J Clin Psychol 43(3):245–265. https://doi.org/10.1348/0144665031752934

108. Clark RA, Pierce AJ, Finn K, Hsu K, Toosley A, Williams L (1998) The impact of alternative approaches to comforting, closeness of relationship, and gender on multiple measures of effectiveness. Commun Stud 49(3):224–239. https://doi.org/10.1080/10510979809368533

109. Hays RD, DiMatteo MR (1987) A short-form measure of loneliness. J Pers Assess 51(1):69–81. https://doi.org/10.1207/s15327752jpa5101

110. Cohen S, Kamarck T, Mermelstein R (1983) A global measure of perceived stress. J Health Soc Behav 24(4):385–396. https://doi.org/10.2307/2136404

111. Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48. https://doi.org/10.18637/JSS.V067.I01

112. Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: tests in linear mixed effects models. J Stat Softw 82(13):1–26. https://doi.org/10.18637/JSS.V082.I13

113. Stower R, Tatarian K, Rudaz D, Chamoux M, Chetouani M, Kappas A (2022) Does what users say match what they do? Comparing self-reported attitudes and behaviours towards a social robot. In: 2022 31st IEEE international conference on robot and human interactive communication (RO-MAN), pp 1429–1434. https://doi.org/10.1109/RO-MAN53752.2022.9900782

114. Fox J, Gambino A (2021) Relationship development with humanoid social robots: applying interpersonal theories to human-robot interaction. Cyberpsychol Behav Soc Netw 24(5):294–299. https://doi.org/10.1089/CYBER.2020.0181

115. Nielsen YA, Pfattheicher S, Keijsers M (2022) Prosocial behavior toward machines. Curr Opin Psychol 43:260–265. https://doi.org/10.1016/J.COPSYC.2021.08.004

116. Axelsson M, Bodala IP, Gunes H (2021) Participatory design of a robotic mental well-being coach. In: 2021 30th IEEE international conference on robot and human interactive communication, RO-MAN 2021, pp 1081–1088. https://doi.org/10.1109/RO-MAN50785.2021.9515356

117. Hawkley LC, Cacioppo JT (2010) Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 40(2):218–227. https://doi.org/10.1007/S12160-010-9210-8

118. Lampraki C, Hoffman A, Roquet A, Jopp DS (2022) Loneliness during COVID-19: development and influencing factors. PLoS ONE 17(3):0265900. https://doi.org/10.1371/JOURNAL.PONE.0265900

119. Pirhonen J, Tiilikainen E, Pekkarinen S, Lemivaara M, Melkas H (2020) Can robots tackle late-life loneliness? Scanning of future opportunities and challenges in assisted living facilities. Futures 124:102640. https://doi.org/10.1016/J.FUTURES.2020.102640

120. Ruggiero A, Mahr D, Odekerken-Schröder G, Spena TR, Mele C (2022) Companion robots for well-being: a review and relational framework. In: Davis M (ed) Research handbook on services management. Edward Elgar Publishing, Cheltenham, pp 309–330

121. Odekerken-Schröder G, Mele C, Russo-Spena T, Mahr D, Ruggiero A (2020) Mitigating loneliness with companion robots in the COVID-19 pandemic and beyond: an integrative framework and research agenda. J Serv Manag 31(6):1149–1162. https://doi.org/10.1108/JOSM-05-2020-0148/FULL/PDF

122. Engelhart K (2021) What robots can-and can't-do for the old and lonely. https://www.newyorker.com/magazine/2021/05/31/what-robots-can-and-cant-do-for-the-old-and-lonely

123. Chang N (2022) New York trials robot companions for 800 elderly people to combat loneliness|Euronews. https://www.euronews.com/next/2022/05/31/new-york-trials-robot-companions-for-800-elderly-people-to-combat-loneliness

124. Spitale M, Gunes H (2022) Affective robotics for wellbeing: a scoping review. In: 2022 10th International conference on affective computing and intelligent interaction workshops and demos (ACIIW), pp 1–8 . https://doi.org/10.1109/ACIIW57231.2022.10085995

125. Bujarski SJ, Mischel E, Dutton C, Steele JS, Cisler J (2015) The elicitation and assessment of emotional responding. In: Sleep and affect: assessment, theory, and clinical implications, pp 91–118. https://doi.org/10.1016/B978-0-12-417188-6.00005-0

126. Laban G, Morrison V, Kappas A, Cross ES (2022) Informal caregivers disclose increasingly more to a social robot over time. In: CHI conference on human factors in computing systems extended abstracts. CHI EA'22, pp 1–7. Association for Computing Machinery, New York. https://doi.org/10.1145/3491101.3519666

127. Gunes H, Broz F, Crawford CS, der Pütten AR-v, Strait M, Riek L (2022) Reproducibility in human-robot interaction: furthering the science of HRI. Curr Robot Rep. https://doi.org/10.1007/s43154-022-00094-5

128. Henschel A, Hortensius R, Cross ES (2020) Social cognition in the age of human-robot interaction. Trends Neurosci 43(6):373–384. https://doi.org/10.1016/j.tins.2020.03.013

129. Powell H, Laban G, George J-N, Cross ES (2022) Is deep learning a valid approach for inferring subjective self-disclosure in human-robot interactions? In: Proceedings of the 2022 ACM/IEEE international conference on human-robot interaction. HRI'22. IEEE Press, Sapporo, pp 991–996. https://doi.org/10.5555/3523760.3523921

130. Abbasi N, Spitale M, Anderson J, Ford T, Jones P, Gunes H (2022) Computational audio modelling for robot-assisted assessment of children's mental wellbeing. In: 14th International conference on social robotics. https://doi.org/10.17863/CAM.89924

131. Laban G, Kappas A, Morrison V, Cross ES (2023) Opening up to social robots: how emotions drive self-disclosure behavior. In: 2023 32nd IEEE International conference on robot and human interactive communication (RO-MAN). IEEE, Busan, pp 1697–1704. https://doi.org/10.1109/RO-MAN57019.2023.10309551

132. Li J (2015) The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum Comput Stud 77:23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

133. Heerink M, Kröse B, Evers V, Wielinga B (2007) Observing conversational expressiveness of elderly users interacting with a robot and screen agent. In: 2007 IEEE 10th international conference on rehabilitation robotics, ICORR'07, pp 751–756. https://doi.org/10.1109/ICORR.2007.4428509

134. Kiesler S, Powers A, Fussell SR, Torrey C (2008) Anthropomorphic interactions with a robot and robot-like agent. Soc Cogn 26(2):169–181. https://doi.org/10.1521/SOCO.2008.26.2.169

135. Honig S, Oron-Gilad T (2020) Comparing laboratory user studies and video-enhanced web surveys for eliciting user gestures in human-robot interactions. In: Companion of the 2020 ACM/IEEE international conference on human-robot interaction. Association for Computing Machinery, New York, pp 248–250. https://doi.org/10.1145/3371382.3378325

136. Gittens CL, Garnes D (2022) Zenbo on zoom: evaluating the human-robot interaction user experience in a video conferencing session. Digest of technical papers—IEEE international conference on consumer electronics, vol 2022-January. https://doi.org/10.1109/ICCE53296.2022.9730259

137. Hoffman G, Zhao X (2020) A primer for conducting experiments in human-robot interaction. J Hum-Robot Interact. https://doi.org/10.1145/3412374

138. Salkind NJ (2010) Within-subjects design. In: Encyclopedia of research design. SAGE Publications, Inc., https://doi.org/10.4135/9781412961288

139. Charness G, Gneezy U, Kuhn MA (2012) Experimental methods: between-subject and within-subject design. J Econ Behav Org 81(1):1–8. https://doi.org/10.1016/J.JEBO.2011.08.009

140. Lepper HS (1998) Use of other-reports to validate subjective well-being measures. Soc Indic Res 44(3):367–379

141. Pavot W, Diener E (1993) The affective and cognitive context of self-reported measures of subjective well-being. Soc Indic Res 28(1):1–20

142. Linton MJ, Dieppe P, Medina-Lara A (2016) Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. BMJ Open 6(7):10641. https://doi.org/10.1136/BMJOPEN-2015-010641

143. Short JC, McKenny AF, Reid SW (2018) More than words? Computer-aided text analysis in organizational behavior and psychology research. Annu Rev Organ Psych Organ Behav 5:415–435. https://doi.org/10.1146/ANNUREV-ORGPSYCH-032117-104622

144. Bak JY, Lin CY, Oh A (2014) Self-disclosure topic model for classifying and analyzing Twitter conversations. In: EMNLP 2014–2014 conference on empirical methods in natural language processing, proceedings of the conference. Association for Computational Linguistics (ACL), pp 1986–1996. https://doi.org/10.3115/V1/D14-1213. https://aclanthology.org/D14-1213

145. Pluta M (2021) Online self-disclosure and social sharing of emotions of women with breast cancer using instagram-qualitative conventional content analysis. Chronic Illn 18(4):834–848. https://doi.org/10.1177/17423953211039778

146. Sermat V, Smyth M (1973) Content analysis of verbal communication in the development of relationship: conditions influencing self-disclosure. J Pers Soc Psychol 26(3):332–346. https://doi.org/10.1037/H0034473

147. Chen L, Hu N, Shu C, Chen X (2019) Adult attachment and self-disclosure on social networking site: a content analysis of Sina Weibo. Person Individ Differ 138:96–105. https://doi.org/10.1016/J.PAID.2018.09.028

148. Law EL-C, Følstad A, Grudin J, Schuller B (2022) Conversational agent as trustworthy autonomous system (trust-CA) (Dagstuhl Seminar 21381). Dagstuhl Rep 11(8):76–114. https://doi.org/10.4230/DagRep.11.8.76

149. Porcheron M, Lee M, Nesset B, Guribye F, van der Goot M, Moore KR, Usbeck R, Paiva A, Pelachaud C, Ruane E, Schuller B, Laban G, Kontogiorgos D, Kraus M, Følstad A (2022) Definition, conceptualisation and measurement of trust. Dagstuhl Rep 11(8):101–105. https://doi.org/10.4230/DAGREP.11.8.76

Guy Laban is a Postdoctoral Research Associate at the Department of Computer Science & Technology of the University of Cambridge, and a member of the Affective Intelligence and Robotics Laboratory (AFAR). Guy pursued his PhD studies in Neuroscience and Psychology as a Marie Skłodowska-Curie Fellow at the School of Psychology and Neuroscience of the University of Glasgow. During his PhD studies, Guy was an Early Stage Research (ESR) member of ENTWINE, the European Training Network on Informal Care, a Marie Skłodowska-Curie Innovation Training Network (ITN) funded by the European Union. Guy's research interests centre on supporting individuals' emotional well-being through interactions with robots. Specifically, Guy investigates how individuals convey their emotions to robots, and how these interactions can be leveraged to support their overall emotional health. Guy studies, develops, and designs robots that facilitate meaningful social exchanges with human users across various contexts. Drawing from principles in computer science, psychology, affective computing, and human-computer interaction, Guy's work aims to create emotionally intelligent and socially supportive robotic systems.

Arvid Kappas is a professor of psychology and Dean at Constructor University. He has been conducting research on emotions for over three decades. Having obtained his Ph.D. at Dartmouth College, NH, USA, he has held positions in Switzerland, Canada, the UK, and Germany. For over two decades, Arvid's research has revolved around the pivotal concepts of Emotion, Cognition, and Social Context. His current research interests focus on various aspects of human emotions, as well as Affective Computing and Social Robotics, aiming to imbue machines with emotional intelligence to enhance our understanding of human behavior in the development of artificial systems like robots or virtual characters. He is a fellow of the Association for Psychological Science and the Society for Personality and Social Psychology. From 2013–2018 he was president of the International Society for Research on Emotion.

Val Morrison is a chartered Psychologist and a full Professor of Health Psychology at the School of Psychology, Bangor University, Wales, UK. Prof. Morrison is also College Lead for Research with Impact at Bangor University. Her research involves prospective studies using mixed methods (quantitative and qualitative) conducted with patients or carers living with a range of chronic health conditions. A key aspect of this work has involved identifying patterns, predictors and consequences of motivations to care and cultural variations therein.

Emily S. Cross received the Ph.D. degree in cognitive neuroscience from Dartmouth College, Hanover, NH, USA. Recently, she joined ETH Zürich to establish the Professorship for Social Brain Sciences in Zurich, Switzerland. Prior to this, she was based jointly with the MARCS Institute, Western Sydney University, Sydney, NSW, Australia; the Department of Cognitive Science, Macquarie University, Sydney; and the School of Psychology and Neuroscience, University of Glasgow, Glasgow, U.K. She leads a vibrant and diverse research team that uses brain imaging techniques, robots, and complex action training paradigms to explore how experience-dependent plasticity and expertise is manifest across brain and behavior, with a particular focus on how our experiences and expectations about artificial agents influence our interactions with these agents. Her work has been supported by a range of funders worldwide, including the European Research Council, the Australian Research Council, Research Councils UK, the Dutch Science Foundation, the Alexander von Humboldt Stiftung, the National Institutes of Health, and the Fulbright Commission.