

Article

Self-Disclosure to a Robot: Only for Those Who Suffer the Most

Yunfei (Euphie) Duan¹, Myung (Ji) Yoon¹, Zhixuan (Edison) Liang² and Johan Ferdinand Hoorn^{1,2,*} 

¹ School of Design, The Hong Kong Polytechnic University, Hong Kong, China; euphie.duan@connect.polyu.hk (Y.D.); ji-myung.yoon@connect.polyu.hk (M.Y.)

² Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China; 18094409g@connect.polyu.hk

* Correspondence: johan.f.hoorn@polyu.edu.hk; Tel.: +528-2766-4509

Abstract: Social robots may become an innovative means to improve the well-being of individuals. Earlier research has shown that people easily self-disclose to a social robot, even in cases where it was unintended by the designers. We report on an experiment considering self-disclosing in a diary journal or to a social robot after negative mood induction. An off-the-shelf robot was complemented with our in-house developed AI chatbot, which could talk about ‘hot topics’ after training it with thousands of entries on a complaint website. We found that people who felt strongly negative after being exposed to shocking video footage benefited the most from talking to our robot, rather than writing down their feelings. For people less affected by the treatment, a confidential robot chat or writing a journal page did not differ significantly. We discuss emotion theory in relation to robotics and possibilities for an application in design (the emoji-enriched ‘talking stress ball’). We also underline the importance of otherwise disregarded outliers in a data set of therapeutic nature.

Keywords: self-disclosure; social robots; diary; emotion theory; relevance; valence



Citation: Duan, Y.; Yoon, M.; Liang, Z.; Hoorn, J.F. Self-Disclosure to a Robot: Only for Those Who Suffer the Most. *Robotics* **2021**, *10*, 98. <https://doi.org/10.3390/robotics10030098>

Academic Editors: Dan Zhang and Xinjun Liu

Received: 12 May 2021
Accepted: 13 July 2021
Published: 29 July 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the outbreak of the COVID-19 pandemic, there has been an upsurge of interest in social isolation, loneliness and depression. People living alone, people with low socio-economic status and, quite unexpectedly, youngsters and students have been found to be at risk of loneliness [1,2]. In the United States, lockdowns and social distancing measures have been associated with increased levels of loneliness, which correlates highly with depression and suicidal ideation. Loneliness remained high, even after distancing measures were relaxed [3]. In the United Kingdom—a country severely impacted by the pandemic—people with COVID-19 likely developed psychiatric disorders and were lonelier, particularly women, adolescents and young adults [4]. In Hong Kong, where our current study took place, COVID-19 even led to “alarming levels of psychiatric symptoms,” with loneliness playing a disadvantageous role [5]. A number of interventions may help to reduce the feeling of loneliness during social isolation, including mindfulness exercises, lessons on friendship, robot pets and programs that facilitate making social contact [6].

Media exposure to negative information, such as war and disasters, may also lead to negative psychological outcomes, particularly feelings of anxiety [7]. It seems that, in developed countries, depression, stress and anxiety have increased among the youth as a result of intensive media use. For example, cases of (attempted) suicide among adolescents have increased since 2010 in the U.S., which may be linked to heavy media usage [8].

1.1. Literature Review

To improve the mental well-being of individuals, a considerable number of studies have focused on the reduction of negative emotions. Emotions are a characteristic human phenomenon [9], which have a huge impact on the lives of individuals, including their judgment and decision making in different contexts which, in turn, influences their mental

well-being [10]. Numerous empirical studies have highlighted that negative emotions are harmful to the health and wellbeing of individuals [10–18].

However, emotions are a complex matter, involving feelings, physiology, cognition, expression and behaviors [9,19]. In appraisal theory [9,20], appraisal (as a component of emotion) is considered both a cognitive [20,21] and an unconscious process [9], which influences emotions through valence judgments.

Valence is regarded, by many psychologists, as a force that attracts individuals to pleasant objects or repels them from unpleasant ones [9,22–24]. Valence has been considered a building block of different emotions [25].

Around the core effect of ‘pleasant’ versus ‘unpleasant’, other aspects of ‘valency’ are the perception of affective qualities, sensitivity to stimuli [23], goal conduciveness, emotion antecedents and consequents [9], coping potential, self-congruency, familiarity, contentment, self-worthiness and moral goodness [24]. Emotions transpire when something is appraised as positive or negative in a certain context; hence, a positive or negative emotion is the result of the valence appraisal process.

Another core concept in emotion theory is the relevance of an event to the goals, needs and concerns of an individual [26]. Relevance is an individual’s response to events appraised as impacting their concerns [9,27,28]. It reflects the personal meaning attached to an event or object. Relevance directs how grave, severe or urgent something is perceived to be. Scherer [28] suggested that relevance can be explained in the sense of “an event having significant and demonstrable bearing on the well-being of the individual”.

Relevance influences the appraisal of an event, which may or may not trigger one’s emotions. As relevance has a significant and demonstrable bearing on the well-being of an individual, stimuli that have such a direct bearing must be salient or of high priority when the event occurs [28]. In our case, self-disclosure to a robot is less relevant to an individual’s concerns when they feel okay than when they feel stressed out.

Another aspect of emotion that has direct bearing in our case is that, in the early stages of appraisal, the novelty of a stimulus also generates emotions [20,28,29]. Novelty is closely related to the perception of affective qualities. Scherer [28] proposed to conceive of novelty in terms of suddenness, unfamiliarity and unpredictability. Yet, novelty does not seem to be as unanimously present as valence and relevance. Novelty does not last forever—it wears out and, at some point, is emotionally not a ‘surprise’ anymore (cf. [30]). In our experiment, thus, novelty is not the focus of interest but, instead, serves as a control variable while examining the effect of different media on the reduction of negative affect.

As our mind and body are closely related, researchers have indicated that negative emotions are associated with bad subjective health and mental well-being [15–18]. One of the popular coping approaches to reduce the level of negative emotions involves emotion disclosure interventions, in which individuals reveal their thoughts and feelings through self-disclosure [31]. Smyth [32] put forward that such disclosure interventions are effective in reducing the level of negative emotions for both males and females. Pennebaker [33] suggested that self-disclosure reduces the psychological work of actively having to inhibit emotions and thoughts about negative events (e.g., trauma), which reduces the associated stress.

To self-disclose, talking with a psychiatrist and journal writing are methods which have been widely adopted in psychotherapy. A variety of studies have examined journal writing to reduce distress [34–37]. Journal writing has beneficial effects, particularly for college students [38]. Writing, as an intervention, can transfer the non-verbal memories into a verbal form that helps to reorganize the memories, resulting in stress reduction [39,40].

The meta-analysis by Frisina, Borod and Lepore [41] found that writing improved health outcomes ($d = 0.19$); however, the effect was stronger for physical outcomes ($d = 0.21$) than for psychological outcomes ($d = 0.07$) (ibid.). In accordance, Pascoe [42] stated on the effectiveness of writing to reduce the level of negative emotions, but the results were limited and need further study. The most beneficial form of writing seems to include large numbers of positive emotion words and a moderate number of negative emotion

words. Participants who used too many or too few negative emotion words benefited less from a writing intervention [42], such that writing may be contra-indicated for individuals with, for instance, alexithymia, who are unable to express emotions [43]. Moreover, the studies conducted by [44–46] pointed out that the physical presence of a therapist is what moderated the negative emotions, rather than the writing itself.

The problem at present is that, worldwide, mental-health workers, therapists and psychiatrists are in short supply [47]. Luckily, however—and quite unexpectedly—since the release of the Rogerian chatbot therapist ELIZA [48], people do not merely share their secrets with fellow humans, but also with their Apple Siri voice agent (see, e.g., [49]), as well as with conversation and companion robots (see, e.g., [50]). Perhaps, then, that social robots may be an ‘AI-in-Design’ alternative to practice emotion-disclosure interventions with—provided that they work well, of course.

In that respect, Wada, Shibata, Saito, Sakamoto and Tanie [51] showed that social robots can alleviate adverse emotions, such as loneliness and stress. As measured on a geriatric-depression scale, as well as a ‘face scale,’ the level of depression of participants significantly decreased after interaction with a social robot (*ibid.*). Jibb, Birnie, Nathan, Beran, Hum, Victor and Stinson [52] found that talking to a robot reduced the level of distress among children who had undergone cancer treatment. Dang and Tapus [53] found that social robots can assist humans during emotion-oriented coping, using a stress-eliciting game played together with a robot. Cabibihan, Javed, Ang and Aljunied [54] provided evidence that robots work well for autistic children, as they can improve their adaptive behaviors (see, e.g., [55–58]) and even invite self-disclosure in adolescents with autism spectrum disorders [59]. Social robots also may increase the mental well-being of older adults, through perceived emotional support and interaction [60].

In psychotherapy, robots may meet the special needs of individuals with cognitive, physical or social disabilities [61]. The meta-analysis conducted by [62] indicated that, in overall robot-enhanced psychotherapy, robots have medium-sized significant effects on the improvement of behavior, but not so much on cognitive and subjective aspects. Yet, individual studies sometimes do show that social robots improve performance on the subjective and cognitive level as well (see, e.g., [51,63,64]).

1.2. Research Question and Hypotheses

In view of the generally positive therapeutic effects of robots in reducing stress and anxiety, our research question is whether social robots can offer an alternative to traditional diary-writing to ‘let off some steam,’ particularly in coming to terms with negative valence emotions after violent-media exposure. We expected that social robots would do better than writing down feelings, as the robot more closely resembles talking to a person (i.e., a virtual therapist), and writing may not be everybody’s preferred means of expression. Therefore, we propose (H1) that a social robot that invites self-disclosure from its user decreases the level of negative emotions more than pencil-and-paper approaches. As a medium (H2), a social robot that invites self-disclosure will be regarded as more relevant to the user’s goals and concerns than pencil-and-paper approaches.

2. Materials and Methods

2.1. Participants and Design

After obtaining approval from the institutional Ethical Review Board (filed under HSEARS20200204003), voluntary participants ($N = 45$; $M_{Age} = 24.9$, $SD_{Age} = 3.29$; 55.6% female, 44.4% male; Chinese nationality) were randomly assigned to a between-subject experiment of self-disclosure in a Robot ($n = 24$; 54.2% female) versus Writing condition ($n = 21$; 57.1% female) after negative-mood induction. All participants had university training at the master level, except for four doctorate degrees, three bachelors and one with a diploma degree. Informed consent was obtained formally from all participants. They did not receive any credits or monetary rewards.

2.2. Procedure

Participants were brought into a dimly lit and shielded-off section of the experimental room and were seated in front of a laptop. The experiment consisted of negative-mood induction, followed by self-disclosure through one of two media, after which participants filled out an online questionnaire in the Qualtrics environment for administration of surveys and experiments (<https://www.qualtrics.com/>, accessed on 12 May 2019).

In the induction part, participants were confronted with a 10 min, 6 s video compilation of three documentaries about a serious earthquake incident that happened in Wenchuan Sichuan, China, in 2008. Viewing negative media, including videos, images and text, can effectively induce negative emotions with an increasing activation of the aversive system [65,66]. In accordance with [67], who concluded that video is the most effective means of mood induction, we prepared a video on the Sichuan earthquake, such that the contents were culturally related to our participants, thus bringing relevance and realness to the experience.

After the video and 30–40 s of instruction, participants either talked to a robot about their experiences during the video, or wrote them down on paper. Neither the robot nor writing utensils were visible before the self-disclosure. The self-disclosure session took about 10 min. The movements of the robot and text input were handled by remote control (Wizard of Oz), and the conversation was handled autonomously by our in-house developed AI chatbot (detailed in the following section).

After the self-disclosure ended, participants filled out a 30-item structured questionnaire (Appendix A), reporting on their assessment of the video clip and either talking to the robot or writing the journal page. Appendix A shows the English translation of the Chinese version in the robot condition. Supplementary Materials S1 provides both questionnaire versions—that is, for the robot and writing—in Chinese and English. The items of the questionnaire were presented in blocks, with pseudo-random sequences of items within blocks, which was different for each participant. We ended the questionnaire with questions inquiring about demographic information. Upon completion, participants were thanked for their participation and debriefed.

2.3. Apparatus and Materials

2.3.1. Video Materials

The video materials for negative-mood induction were 10 min and 6 s long and were composed of video excerpts from the following three Sichuan earthquake Internet documentaries:

Internet video in memory of the Wenchuan Sichuan earthquake 10th anniversary (cut from 00:02–01:19). Available from <https://www.bilibili.com/video/av23087386/> (accessed on 13 June 2019)

Dazzz2009 (31 December 2008). Internet video record of 512 earthquake in Dujiangyan (cut from 01:20–01:59). Available from https://www.youtube.com/watch?v=Vz0nGbl81fM&list=PLf2PpWDjx1d6rVUW0vaGFzhvIr_nRo_8&index=2 (accessed on 13 June 2019)

Lantian777 (16 May 2008). Internet video 10 min after Wenchuan Sichuan earthquake (in full). Available from <https://www.youtube.com/watch?v=PI5KL7nvU28> (accessed on 14 June 2019)

2.3.2. Robot Embodiment

For a person to self-disclose, there should be a certain level of trust in the conversation partner, which is also true in human–robot interactions [49,68]. Therefore, we looked for a small, non-threatening toy-like robot that, in appearance, would stay far from uncanny effects (cf. [69]). The chatbot part was trained such that superfluous faults in responding were kept to a minimum (cf. [70]).

The robot of our choice was a Robotis DARwIn Mini: a 27 cm tall 3D printable, programmable and customizable miniature humanoid robot which can connect to a laptop through Bluetooth. The robot could stand up and move its arms while speaking through

an AI chatbot. Technical details for the DARwIn Mini can be found in Supplementary Materials S1. The actions that DARwIn could execute during the experiment, such as waving and raising its arms, were controlled remotely.

2.3.3. Self-Disclosure Chatbot

The DARwIn Mini could not speak; therefore, we created our own chatbot, using DARwIn Mini as the humanoid embodiment of our self-disclosure inviting AI chatbot. Next, we provide a concise account of the development of both the hardware and software. Supplementary Materials S1 offers further specifications.

Hardware development. Two main components comprised the hardware of our chatbot: the core board, a Raspberry Pi Zero (WH) and an extension board, which was connected to the speaker. These two boards were engineered into an integrated circuit. Figure 1 offers an impression of the hardware prototype chatbot.

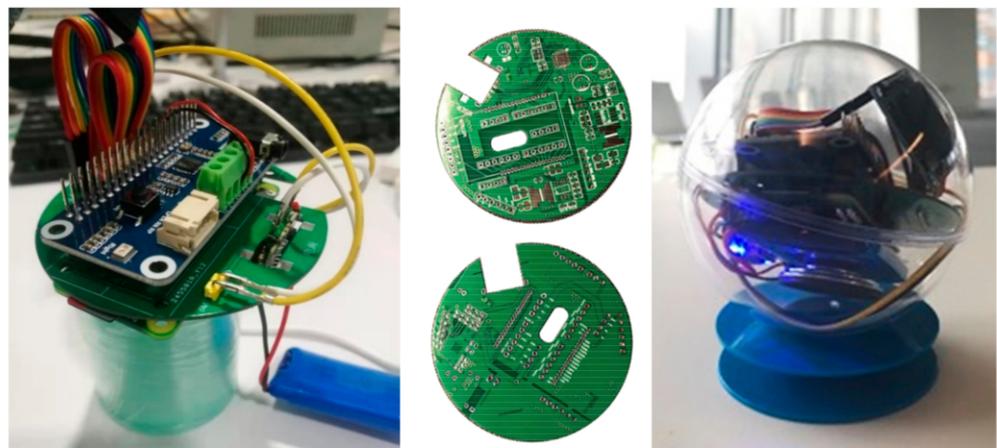


Figure 1. Hardware prototype of the chatbot with in-house-engineered integrated board.

Software development. To create a chatbot adjacent to the DARwIn Mini, we set up a homepage for test subjects to assess the chatbot system (for details on the chatbot, please refer to Supplementary Materials S1; www.roboticmeme.com (accessed on 25 June 2019)). For website development, we used Semantic UI as the front-end framework (<https://semantic-ui.com/> (accessed on 9 May 2019)) and Node.js as the back-end (<https://nodejs.org/en/> (accessed on 9 May 2019)). We tentatively called the chatbot MEME and invited test subjects to share their secrets with MEME in our test environment. The chatbot on the website had speech recognition in Putonghua, Cantonese and English, using a Turing robot API. To increase the traffic on our website, we also created an official WeChat account and used Python to run a server in Google Cloud (<https://cloud.google.com/> (accessed on 11 May 2019)). On WeChat, we used Chill chat with the Xiaohuangji corpus for information retrieval.

We designed a hierarchical chatting system, consisting of three layers: (1) A rule-based layer focused on certain specific chatting tasks (Eliza.py and regular expressions); (2) an information retrieval system that searched the answer from a corpus built from Weibo conversations and conversations about movies; and (3) a generation layer that used the general-purpose encoder seq2seq, as well as a Generative Adversarial Network—a machine-learning tool—to generate a response (<https://github.com/google/seq2seq>; https://en.wikipedia.org/wiki/Generative_adversarial_network (accessed on 11 May 2019)). We adopted the *k*-means algorithm for sentence vector clustering. After many iterations of improvement, the final model could effectively answer a question.

For natural language understanding, we installed a Rasa stack and, so, made the conversation somewhat more contextualized (<https://rasa.com/> (accessed on 11 May 2019)). For Rasa to estimate what a user means to say, we classified a number of conversational topics that had to do with negative experiences. Therefore, we analyzed the contents of

a complaining website and ran a spider program to capture the comments of users, after which we carried out data mining for hot topics.

For training, we sampled a 2-year record of almost 500 pages and nearly 10,000 comments. Then, we tokenized these utterances and identified the high-frequency items ('hot topics').

An impression of the results is depicted in Figure 2.

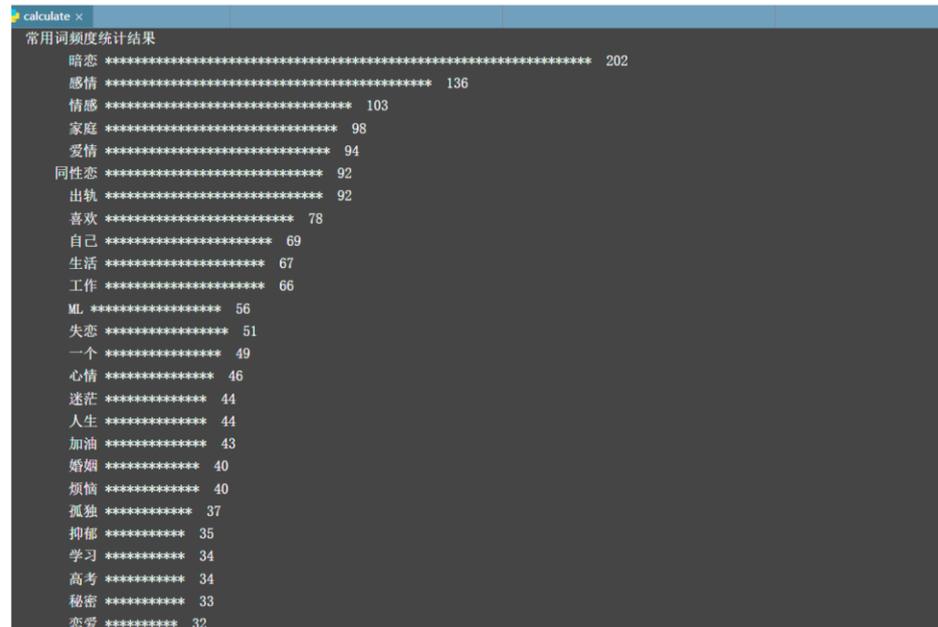


Figure 2. Frequency statistics for hot topics that people complained about.

People worried most about:

(unrequited love)

(sentience)

(affect)

(family)

(homosexuality)

(infidelity)

(love crush)

(self)

(life)

(work)

(ML: making love/sex)

(breaking up (in a romantic relationship))

(being alone)

(mood)

(get lost)

(life)

(to cheer up)

(marriage)

(trouble and worry)

(loneliness)

(depression)

(study)

(Nationwide Unified Examination for Admissions to General Universities and Colleges)

(secrets)

(love relationships)

The complete set-up of the self-disclosure AI chatbot is shown in Figure 3. The sing, movie, poem and weather options were not used in the actual experiment.

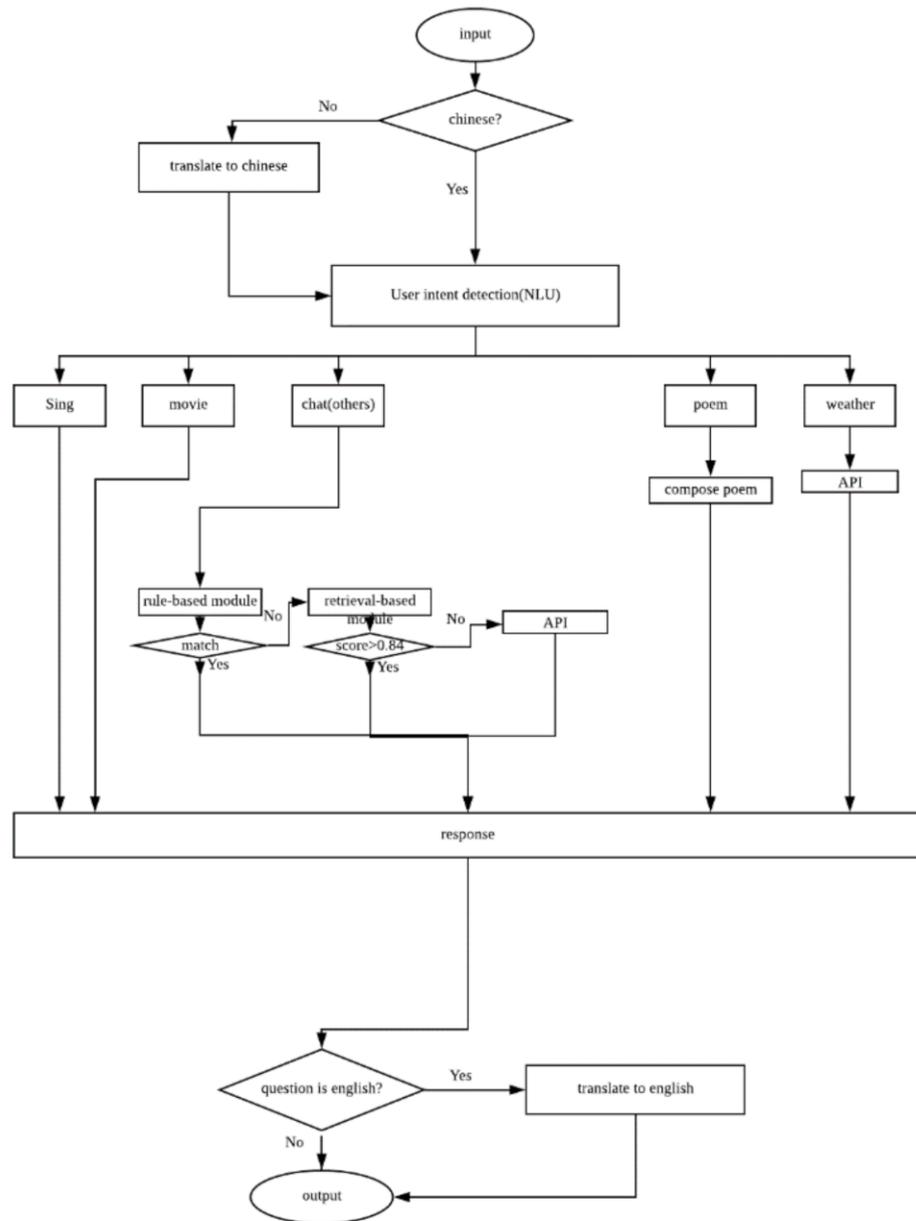


Figure 3. Flowchart for our self-disclosure AI chatbot.

For the experiment, we installed our chatbot system in a voice kit that stood behind the DARwIn Mini. We did not install voice-recognition software, due to its inefficiency (i.e., it is slow and inaccurate). Therefore, we employed a partial Wizard of Oz (WOz) set-up as state-of-the-art voice-recognition, particularly in four-tone Putonghua and, worse, nine-tone Cantonese, is still in its infancy, which would have disturbed the therapeutic effect. Listening, then, was performed in WOz mode, whereas responding was carried out fully autonomously by the system: An experimenter not visible to the participant input the participant’s responses, while information processing and replying to the participants was carried out autonomously by our AI. Figure 4 exhibits the interaction flow.

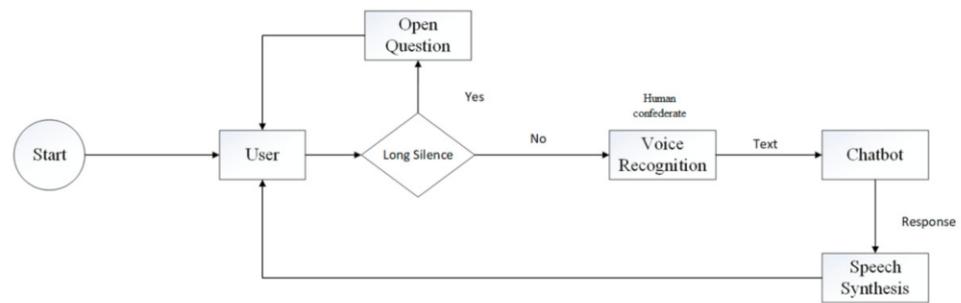


Figure 4. Human–robot interaction flowchart.

Together, the DARwIn Mini, standing in front of the voice kit carrying our self-disclosure AI chatbot, formed the ‘robot condition’ in our experiment. Figure 5 shows the final set-up.



Figure 5. DARwIn Mini placed in front of the voice kit with self-disclosure AI chatbot.

We constructed the conversation following psychotherapeutic guidelines (see [71]). For example, open questions, such as “How do you feel about that?” were asked, in order to guide the participants’ reflections on their experience. During the conversation, only minimal encouragement (e.g., “Yes, I see”) was provided by the robot. The open questions that were coded into the chatbot were also posed to participants in the writing condition, during their instruction. In inviting self-disclosure, the robot basically followed social norms from social penetration theory [72]. Based on [73], however, the robot did not share secrets with its user and did not (need to) apply reciprocity, although this is an important social rule in human interaction (cf. Psychopathology Committee of the Group for the

Advancement of Psychiatry [74]). Thus, the robot was not self-disclosing but invited self-disclosure by asking open questions [50].

Note that DARwIn Mini performed movements and that the participants could touch it, but emotions were expressed only through speech. The interaction possibilities were limited to speaking to the robot, similar to how one can only write on paper.

2.4. Measures

For measurement, we worked from a dimensional model of valence and relevance (cf. [75]), rather than a categorical model, which classifies emotions by name ('sad' or 'happy'). Emotion words are fuzzy [19] and appraisal of an event may elicit a variety of emotions [9,21]. Different people interpret events differently and, so, different emotions are generated after the same event. If negative emotions are presented only by name, then consensus among the participants may be low. As appraisal is a dynamic process [9,76,77], the possibility exists that an individual experiences multiple emotions due to a single event. It is difficult to list all possible negative emotions in a questionnaire by name and not measure fatigue effects eventually. Therefore, we assessed the core concepts of valence (positive/pleasant vs. negative/unpleasant), providing a more fundamental process, compared to aspects of valence that require associative (and, sometimes, conceptual) processing [28].

In self-reported instruments such as Positive and Negative Affect Schedule (PANAS), valence is conceived of in two unipolar dimensions [78]. Each affective direction would be mediated by an independent neural pathway (see, e.g., [79]). This is in contrast to earlier approaches that maintain a bipolar measurement (for a discussion, see [80,81]). Moreover, two unipolar scales ($0 \rightarrow n$, $0 \rightarrow p$) can, at one point, stand in a bipolar constellation ($n \leftrightarrow p$); however, from a bipolar measurement, one can never return to a unipolar conception. Therefore, we decided on two unipolar dimensions to measure affect and constructed a structured questionnaire with more items on a measurement scale, featuring positive (indicative) as well as negative (counter-indicative) items. This approach also remedied potential answering tendencies.

We used two versions of a structured questionnaire, appropriate to one of the two conditions: Talking with the robot or journal writing on a piece of paper (Appendix A and Supplementary Materials S1). The questionnaire was constructed with respect to the emotion literature (e.g., [9,23,28]) and ran four measurement scales: Valence after the movie but before treatment (robot or writing), Valence after treatment, Relevance and Novelty. Together with the Demographic information, Novelty served as a control.

Items were Likert-type statements following a 6-point rating scale (1 = strongly disagree, 6 = strongly agree). One half of the items on each measurement scale consisted of four indicative statements, and the other half consisted of counter-indications. Blocks of related items were offered in pseudo-random order, differing for each participant. Items within blocks also were pseudo-randomly presented to each participant.

The measurement scale 'Valence before treatment' (*ValB*) consisted of four indicative items—for example, "I feel good"—and four counter-indicative items—for example, "I feel bad." We used the same items for the measurement of Valence after talking to the robot or writing on paper, but adjusted the wording to the situation. Thus, 'Valence after treatment' (*ValA*) also had four indicative and four counter-indicative items. Relevance of robot or writing to goals and concerns (i.e., personal emotion regulation) was measured through two indicative items (e.g., '... is useful') and two counter-indicative items (e.g., '... is worthless').

To control for a possible confounding effect of the robot as a novel means to regulate emotions, the Novelty scale was composed of three indicative items (e.g., '... is new') and three counter-indicative items (e.g., '... is commonplace').

The collected demographics included information about the participant's Gender, Age, Education level and Country. At the end of the questionnaire, participants could leave their comments.

Then, we conducted reliability analysis on our measurement scales (for elaboration, see Supplementary Materials S1). For the variables of theoretical interest, all measurement scales, with all items included, achieved good to very good reliability in the first run (Cronbach's $\alpha \geq 0.82$). This was so for the separate sub-scales of Valence (4 items each) and for their combinations (*ValB* and *ValA*, 8 items each), as well as for Relevance (4 items). After repair, the control variable of Novelty (5 items) had Cronbach's $\alpha = 0.77$.

To test the discriminant validity, we performed Principal Component Analysis with Varimax rotation on Valence-after (*ValA*), Relevance and Novelty. Indicative items formed a positive-Valence sub-scale, as the counter-indicative items clustered into a negative-Valence sub-scale. Items on the Relevance scale neatly fell in line, as intended. Novelty showed some spread over both Valence and Relevance; however, because it was a control variable, we kept the scale intact and observed its tendency to coalesce with variables of theoretical interest (as detailed in the Results section).

Then, we calculated the means (*M*) across the items on a scale and performed an outlier analysis for Valence (before and after), Relevance and Novelty. We found that participant 9 was an outlier in *MValB*, and participant 39 in *MValA*. Participants 5 and 21 were outliers for *MValAi*. Participants 39, 27, 38 and 33 were outliers in *MValAc* (see Supplementary Materials S1). There were no outliers in *MNov*, *MRel*, *MValBc* and *MValBi*. We performed our effects analyses with and without these outliers.

3. Results

3.1. Demographics

We checked the countries that participants came from, but only participant 31 reported that she was from Africa; the rest were from China. Inspection of the scatter plot, however, showed that number 31 was not in the zone of outliers. Therefore, we decided to treat this person as one in the same sample and did not treat her differently in the analyses.

Next, we checked whether Age was correlated with the eight dependent variables: Valence-bipolar (before and after), Positive and Negative Valence (both before and after), Relevance and Novelty. We calculated Pearson bivariate correlations (two-tailed) and found no significant relations. Age did have a near-significant weak negative correlation with Positive Valence-before ($r = -0.27$, $sig. = 0.08$), indicating that, with higher age, people were less positive after viewing the earthquake video.

Then, we examined whether Gender had an influence on the eight dependent variables. We ran a MANOVA (Pillai's Trace) to check the effect of Gender, but found no significant multivariate effects ($V = 0.11$, $F_{(7,37)} = 0.68$, $p = 0.688$); however, Gender did exact a small univariate effect on the experience of Novelty ($F_{(1,41)} = 4.18$, $p = 0.047$, $\eta_p^2 = 0.09$). Throughout, females experienced more Novelty ($M = 4.03$, $SD = 0.83$) than males ($M = 3.50$, $SD = 0.87$). However, Novelty was a control variable in our experiment, and was not considered to be of theoretical interest. Therefore, we concluded that Gender did not have a significant effect on the variables theoretically related to our hypotheses.

Among all participants, there were four with doctorate degrees, three with bachelor's degrees and one with a diploma degree. The rest all had master's degrees. We found participant 39, with a doctorate degree, to be one of the outliers to the scale means. Thus, we excluded this participant from the effect analysis of Educational background.

We put the seven participants with a degree other than master's in one group and randomly chose seven other participants (who were not outliers) with a master's degree in the other group. We performed an independent samples *t*-test, in order to check whether Education had effect on the eight dependent variables related to our theoretical hypotheses. We ran this test five times, each time with a different set of masters graduates, and found that Education did have an effect on some of the theoretical variables in certain group comparisons (see Supplementary Materials S1). Therefore, we constructed two data sets, one with all 45 participants (24 in the robot group and 21 in the writing group) and the other with 31 participants (17 in the robot group and 14 in the writing group), excluding

the outliers and the participants with a non-master's degree as educational background. These separate sets were used to assess our hypotheses.

3.2. Manipulation Check: Emotional Effects after Negative-Mood Induction and after Treatment

We wanted to control whether any emotion at all was provoked by the shocking video footage of the earthquake, and whether the treatment (robot or writing) evoked any change in emotion. Or, did everything remain at a scale value of 1 (no emotions reported)?

For $N = 45$ and $n = 31$, we ran a one-sample t -test (two-tailed) with 1 as the test value, in order to see if any negative (or positive) emotions occurred after mood induction, as well as after treatment (Table 1).

Table 1. One-sample t -tests (two-tailed, 1 is the test value), checking whether emotions occurred at all.

Variables	Mood Induction		
	t	p	n
$MValBi$	8.67	0.00001	45
$MValBc$	16.44	0.00001	45
$MValBi$	7.00	0.00001	31
$MValBc$	15.38	0.00001	31
Variables	Treatment		
	t	p	n
$MValAi$	17.83	0.00001	45
$MValAc$	10.35	0.00001	45
$MValAi$	18.65	0.00001	31
$MValAc$	9.39	0.00001	31

For both $N = 45$ and $n = 31$, after the earthquake clips (Table 1, Mood induction), more negative than positive mood was induced, as intended. For both $N = 45$ and $n = 31$, after Treatment (robot or writing), more positive than negative emotions were felt after either talking to a robot or writing a diary page, as intended.

To check whether before–after effects of the treatment actually occurred, we also ran paired-samples t -tests (two-tailed) in both data sets (i.e., $N = 45$ and $n = 31$). Note that these were not tests of our hypotheses, but a mere inspection whether any affective shifts happened at all (Table 2).

Table 2. Paired-samples t -tests (two-tailed) for treatment effects on Valence.

Variables	Before–After Treatment		
	T	p	n
$MValBc-MValAc$	9.34	0.00001	45
$MValBc-MValAc$	9.42	0.00001	31
$MValBi-MValAi$	−7.16	0.00001	45
$MValBi-MValAi$	−7.24	0.00001	31

From Table 2, we may conclude that participants, after the treatment, became less negative (i.e., $MValBc$ was significantly larger than $MValAc$); in addition, after treatment, they became more positive (i.e., $MValBi$ was significantly smaller than $MValAi$). Whether through a robot or through writing, the treatment had an effect on the expected direction and, so, the manipulation worked.

3.3. Effect of Media (Robot vs. Writing) on Valence and Relevance

To analyze the changes in Valence after talking to a robot or writing a diary page, we computed three mean difference scores:

- Valence bipolar: $\Delta Val = MValA - MValB$;
- Positive Valence: $\Delta ValP = MValAi - MValBi$;
- Negative Valence: $\Delta ValN = MValAc - MValBc$.

In Table 3, ΔVal , $\Delta ValP$, $\Delta ValN$, $MRel$ and $MNov$ are shown for the two conditions (robot vs. writing). The top half of Table 3 shows the averages for the entire sample ($N = 45$), the bottom half shows those with the suspected cases excluded ($n = 31$).

Table 3. Valence, Relevance and Novelty for robot and writing.

Variables	Robot			Writing		
	Mean	SD	n	Mean	SD	n
ΔVal	1.77	1.26	24	1.11	0.81	21
$\Delta ValP$	1.75	1.31	24	0.89	1.06	21
$\Delta ValN$	1.78	1.30	24	1.32	0.84	21
$MRel$	4.19	0.99	24	3.98	1.33	21
$MNov$	4.10	0.86	24	3.42	0.77	21
$N = 45$						
Variables	Robot			Writing		
	Mean	SD	n	Mean	SD	n
ΔVal	1.98	1.11	17	1.33	0.83	14
$\Delta ValP$	1.99	1.08	17	1.05	1.17	14
$\Delta ValN$	1.97	1.27	17	1.61	0.76	14
$MRel$	4.35	0.96	17	4.27	1.08	14
$MNov$	4.13	0.95	17	3.53	0.78	14
$n = 31$						

3.3.1. Effects on Bipolar Valence and Relevance

Next, we performed a General Linear Model (GLM) Multivariate analysis of Media (2: robot vs. writing) on ΔVal and $MRel$ (grand mean scores), with $MNov$ as a covariate (Table 4). We did this for $N = 45$ and $n = 31$ separately. For an extensive report, see Supplementary Materials S1.

For the full data set ($N = 45$), with Novelty as a covariate, we did not find significant multivariate effects (Table 4, first row); however, we did find multivariate effects for $MNov$, which covaried quite strongly with $MRel$.

With Novelty excluded from the analysis, the pattern of multivariate effects was similar as before (Table 4, fourth row). Officially, we should stop our scrutiny here. Yet, when we looked into the main effect of Media on ΔVal , we observed that, without Novelty, the effect became significant. As a trend, beneath the surface, it seemed that talking to a robot ($M_{\Delta Val} = 1.76$, $SD = 1.25$) had a more positive impact on Valence (bipolar conception) than writing ($M_{\Delta Val} = 1.10$, $SD = 0.81$) after negative mood induction.

For the reduced data set ($n = 31$), with Novelty as a covariate, Media (robot vs. writing) did not exert any significant multivariate effects on ΔVal or $MRel$ (Table 4, bottom). Novelty ($MNov$) covaried with other variables, but this was significant for $MRel$ alone. With Novelty discarded in the analysis, the pattern of results did not change. Without the outliers, any positive change in valence caused by either robots or writing remained absent.

Table 4. Multivariate effects of Media (robot vs. writing) on the grand mean scores of $\Delta Valence$ and $Relevance$, with $Novelty$ as a covariate.

Robot vs. Writing on:						
	<i>V</i>	<i>F</i>	<i>df</i> _{1,2}	<i>p</i>	η_p^2	<i>N</i>
ΔVal and $MRel$ with $MNov$	0.09	1.98	2,41	0.151	0.09	45
($MRel$ with) $MNov$	0.39	12.92	2,41	0.000	0.39	45
$MNov$		25.91	1,42	0.000	0.38	45
ΔVal and $MRel$	0.09	2.09	2,41	0.136	0.09	45
ΔVal		4.23	1,43	0.046	0.09	45
Robot vs. Writing on:						
	<i>V</i>	<i>F</i>	<i>df</i> _{1,2}	<i>P</i>	η_p^2	<i>n</i>
ΔVal and $MRel$ with $MNov$	0.09	1.32	2,27	0.285	0.09	31
($MRel$ with) $MNov$	0.38	8.33	2,27	0.002	0.37	31
$MNov$		15.40	1,28	0.001	0.36	31

3.3.2. Effects on Positive Valence, Negative Valence and Relevance

For $N = 45$, we ran two GLM Repeated measures of Media (two conditions) on within-subject factor ($\Delta ValP$ vs. $\Delta ValN$), with $MRel$ and $MNov$ separately as covariates (Table 5). We found no significant multivariate effects on unipolar ($\Delta ValP$ vs. $\Delta ValN$): not for the interaction with Media, not with $MRel$ as covariate, and not with $MNov$ as covariate.

Table 5. Repeated measures effects of Media (robot vs. writing) on the grand mean scores of $\Delta Valence$ (positive vs. negative) with $Relevance$ and $Novelty$ as separate covariates.

Robot vs. Writing on:						
	<i>V</i>	<i>F</i>	<i>df</i> _{1,2}	<i>p</i>	η_p^2	<i>N</i>
$\Delta ValP$ vs. $\Delta ValN$	0.05	2.02	1,42	0.162	0.05	45
$\Delta ValP$ vs. $\Delta ValN$ with $MRel$	0.02	0.71	1,42	0.406	0.02	45
$\Delta ValP$ vs. $\Delta ValN$ with $MNov$	0.00	0.004	1,42	0.951	0.00	45
$\Delta ValP \cup \Delta ValN$ with $MRel$		3.79	1,42	0.058	0.08	45
$\Delta ValP \cup \Delta ValN$ with $MNov$		2.04	1,42	0.161	0.05	45
$\Delta ValP \cup \Delta ValN$		4.23	1,43	0.046	0.09	45
Robot vs. Writing on:						
	<i>V</i>	<i>F</i>	<i>df</i> _{1,2}	<i>p</i>	η_p^2	<i>n</i>
$\Delta ValP$ vs. $\Delta ValN$	0.09	2.63	1,28	0.116	0.09	31
$\Delta ValP$ vs. $\Delta ValN$ with $MRel$	0.01	0.30	1,28	0.588	0.01	31
$\Delta ValP$ vs. $\Delta ValN$ with $MNov$	0.004	0.13	1,28	0.725	0.004	31
$\Delta ValP \cup \Delta ValN$		3.14	1,28	0.087	0.10	31

With $MRel$ included, we did find a significant main effect ($p = 0.046$) of Media across $\Delta ValP$ and $\Delta ValN$ (non-unipolar Valence), as shown in Table 5. With $MNov$ included, however, this main effect was not significant any more ($p = 0.087$). This pattern of results remained the same without the covariates, except that, as before, the effect of Media across $\Delta ValP$ and $\Delta ValN$ (non-unipolar Valence) became significant (Table 5, row 6).

For $n = 31$ (Table 5, bottom), we again ran two GLM Repeated measures of Media (2 conditions) on ($\Delta ValP$ vs. $\Delta ValN$), with $MRel$ and $MNov$ as a separate covariate, respectively. As before, we found no significant multivariate effects on ($\Delta ValP$ vs. $\Delta ValN$): not for

the interaction with Media, not with *MRel* as covariate and not with *MNov* as covariate. Without the emotional outliers, the main effect of Media on the unipolar conception of Valence ($\Delta ValP$ and $\Delta ValN$) remained absent (Table 5, bottom row). Without the covariates, the pattern of these results did not change.

Overall, we saw that the only small significant effect we could establish for the theoretical variables was with $N = 45$, without *MNov* as a covariate, in a bipolar conception of Valence (ΔVal). We wondered, then, how this could be the case, as the mood induction and the treatment had been so successful, according to the *t*-test (Section 3.2).

3.4. Effect of Media on Valence and Relevance for Those Who Felt Most Negative

In clinical trials, it is good practice to contrast a control group with a treatment group, in order to measure the effects of a drug or medical device (see, e.g., [82] p. 2). We attempted the same, but with depressed people (after mood induction) and using two different media (robot vs. pen-and-paper). However, another approach in clinical research is to try a drug on healthy volunteers versus patient volunteers; this is what we, so far, failed to recognize: part of the participants may not have been affected much by the mood induction and, therefore, did not need treatment or comfort from our robot or journal writing. After all, they were not distressed—they did feel the emotion but were ‘immune to the affliction’ and, so, the treatment was superfluous (i.e., a sub-sample ceiling effect).

Therefore, we performed a median split for both data sets $N = 45$ and $n = 31$ on the variable *MValBc* (Negative Valence before treatment). In the data set with $N = 45$, with the outliers included, 23 participants were on the side of feeling most negative. Of these, 12 were in the robot condition and 11 were in the writing condition.

For $n = 31$, without the outliers, 17 participants felt most negative, 10 of which talked to a robot after viewing the footage and 7 carried out the writing task. Table 6 provides the means and SDs for ΔVal , $\Delta ValP$, $\Delta ValN$, *MRel* and *MNov* for talking to a robot or writing a journal page for those participants who felt very negative after watching the earthquake video.

Table 6. Valence, Relevance and Novelty of the most negatively affected participants in robot and writing conditions ($n = 40$).

Variables	Robot			Writing		
	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>
ΔVal	2.74	0.83	12	1.56	0.84	11
$\Delta ValP$	2.68	0.84	12	1.31	1.16	11
$\Delta ValN$	2.79	0.96	12	1.77	0.75	11
<i>MRel</i>	4.17	1.04	12	4.25	1.31	11
<i>MNov</i>	3.27	0.92	12	4.52	0.56	11
With emotional outliers: $n = 23$						
Variables	Robot			Writing		
	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>
ΔVal	2.65	0.80	10	1.69	0.83	7
$\Delta ValP$	2.55	0.81	10	1.42	1.21	7
$\Delta ValN$	2.75	0.95	10	1.96	0.78	7
<i>MRel</i>	4.13	0.80	10	1.70	0.83	7
<i>MNov</i>	3.45	1.02	10	4.49	0.64	7
Without emotional outliers: $n = 17$						

3.4.1. Valence as a Bipolar Scale in High-Negative Subjects

For $n = 23$, GLM Multivariate on ΔVal and $MRel$ showed that, with Novelty ($MNov$) as a covariate, Media (robot vs. writing) exerted significant multivariate effects (Table 7). Media had a significant and moderately strong univariate effect on ΔVal , but not on $MRel$.

Table 7. Multivariate effects of Media (robot vs. writing) on the grand mean scores of $\Delta Valence$ and $Relevance$ with $Novelty$ as covariate for highly negative subjects.

Robot vs. Writing on:						
	V	F	$df_{1,2}$	P	η_p^2	n (outliers included)
ΔVal and $MRel$ with $MNov$	0.46	8.09	2,19	0.003	0.46	23
ΔVal		8.80	1,20	0.008	0.31	23
$MRel$		2.16	1,20	0.160	0.10	23
($MRel$ with) $MNov$	0.47	8.42	2,19	0.002	0.47	23
ΔVal with $MNov$		<1	2,19	0.459		23
ΔVal and $MRel$	0.40	6.79	2,20	0.006	0.40	23
ΔVal		11.51	1,21	0.003	0.35	23
$MRel$		0.03	1,21	0.867	0.001	23
Robot vs. Writing on:						
	V	F	$df_{1,2}$	P	η_p^2	n (outliers excluded)
ΔVal and $MRel$ with $MNov$	0.38	3.94	2,13	0.046	0.38	17
ΔVal		4.07	1,14	0.063	0.23	17
$MRel$		2.23	1,14	0.157	0.14	17
($MRel$ with) $MNov$	0.44	5.16	2,13	0.022	0.44	17
$MNov$		10.87	1,14	0.005	0.44	17
(ΔVal with) $MNov$		0.15	1,14	0.700	0.01	17
ΔVal and $MRel$	0.30	3.04	2,14	0.080	0.30	17
ΔVal		5.64	1,15	0.031	0.27	17
$MRel$		0.074	1,15	0.790	0.005	17

$MNov$ also showed significant multivariate effects, but on $MRel$ alone, not on ΔVal . Novelty made things more relevant.

After removing $MNov$ as a covariate, we found that Media still evoked significant multivariate effects, substantiated by a significant and moderately strong effect of Media on ΔVal . There was no significant effect on $MRel$ (Table 7).

With emotional outliers included, then, talking to a robot ($M_{\Delta Val} = 2.74$, $SD = 0.83$) had a more positive impact on Valence (bipolar conception) than writing ($M_{\Delta Val} = 1.56$, $SD = 0.84$) after negative mood induction. The level of novelty of the medium made things more relevant, but neither medium was significantly more relevant than the other. Novelty did not significantly influence the valence result.

For $n = 17$, without outliers (Table 7, bottom), GLM Multivariate on ΔVal and $MRel$ showed that, with Novelty as a covariate, significant multivariate effects were established. There was a main effect of Media on ΔVal , which went in the direction of being significant; however, this was not so for $MRel$.

Multivariate effects for $MNov$ were significant, again when covarying with $MRel$ but not with ΔVal .

After removing $MNov$ as a covariate, we found that no significant multivariate effects were present any more, although ‘under the surface,’ the between-subject effects showed a

significant effect of Media on ΔVal in the expected direction (Table 7): Robot ($M_{\Delta Val} = 2.65$, $SD = 0.80$) was higher than Writing ($M_{\Delta Val} = 1.69$, $SD = 0.83$). There was still no significant effect of Media on $MRel$.

It seemed, then, with the outliers dismissed from the data (and less power due to fewer subjects), that the effects tended to disappear. It was for those who suffered the most that the robot was the most helpful. The novelty aspect of talking to a robot may make the medium more relevant to personal goals and concerns, but is not (or for the less affected only marginally) influential for feeling more positive after a chat with a robot about negative experiences.

3.4.2. Positive and Negative Valence as Two Unipolar Scales in Highly Negative Subjects

For $n = 23$, we ran two GLM Repeated measures of Media (2 conditions) on the within-subject factor ($\Delta ValP$ vs. $\Delta ValN$) with $MRel$ and $MNov$ separately as covariates. Multivariate tests showed that no significant effects occurred for $\Delta ValP$ vs. $\Delta ValN$. The height of positive and negative valences did not differ. The interaction of ($\Delta ValP$ vs. $\Delta ValN$) with Media also was not significant (Table 8), nor was $MRel$ as a covariate. However, the main effect of Media was significant, showing that robots exerted higher levels of undifferentiated Valence (non-unipolar) than writing on paper. We repeated the test with Novelty as the covariate, but $MNov$ did not significantly contribute to any of the effects.

Table 8. Repeated measures effects of Media (robot vs. writing) on the grand mean scores of $\Delta Valence$ (positive vs. negative), with *Relevance* and *Novelty* as separate covariates for highly negative subjects.

Robot vs. Writing on:						
	<i>V</i>	<i>F</i>	<i>df</i> _{1,2}	<i>p</i>	η_p^2	<i>n</i>
$\Delta ValP$ vs. $\Delta ValN$	0.04	0.78	1,20	0.387	0.04	23
$\Delta ValP$ vs. $\Delta ValN$ with <i>MRel</i>	0.003	0.06	1,20	0.815	0.003	23
$\Delta ValP \cup \Delta ValN$		13.54	1,20	0.001	0.40	23
Robot vs. Writing on:						
	<i>V</i>	<i>F</i>	<i>df</i> _{1,2}	<i>p</i>	η_p^2	<i>n</i>
$\Delta ValP$ vs. $\Delta ValN$	0.03	0.48	1,14	0.498	0.033	17
$\Delta ValP$ vs. $\Delta ValN$ with <i>MRel</i>	0.00	0.06	1,14	0.936	0.000	17
$\Delta ValP \cup \Delta ValN$		5.98	1,14	0.028	0.30	17
$\Delta ValP$ vs. $\Delta ValN$ with <i>MNov</i>	0.011	0.16	1,14	0.695	0.011	17
$\Delta ValP \cup \Delta ValN$		4.07	1,14	0.063	0.23	17

Then, we carried out the same test for the $n = 17$ data set. We ran two GLM Repeated measures of Media (2 conditions) on the within-subject factor ($\Delta ValP$ vs. $\Delta ValN$), with $MRel$ and $MNov$ as separate covariates. Multivariate tests showed that no significant effects were obtained for $\Delta ValP$ vs. $\Delta ValN$. Here, again, the height of positive and negative valences did not differ. The interaction of ($\Delta ValP$ vs. $\Delta ValN$) with Media was also not significant, nor was $MRel$ as a covariate (Table 8, bottom part). Yet, the main effect of Media remained significant. Repeating the analysis with Novelty as the covariate did not change these results, except for the main effect of Media, which now merely moved into the direction of being significant.

Thus, with outliers excluded, robots still exerted higher levels of undifferentiated Valence (non-unipolar) than writing on paper. With Novelty included, these positive effects became somewhat more pronounced for the less affected.

3.4.3. Exploratory Analyses

Above, we saw that Novelty mainly affected Relevance, indicating that a medium becomes more relevant when it is newer to those who are emotionally affected, but not too

much. In Section 3.1, we found, in turn, that Novelty was affected by Gender. Therefore, we explored the Media \times Gender effects on Novelty with Univariate ANOVA for both data sets ($N = 45$ and $n = 31$). The research question was whether robots were newer to females than to men, or vice versa?

With $N = 45$, only the main effects were significant: Robots ($M = 4.10$, $SD = 0.87$) were perceived as newer than writing ($M = 3.41$, $SD = 0.77$; $F_{(1,41)} = 9.50$, $p = 0.004$, $\eta_p^2 = 0.19$), which was independent of Gender. Females ($n = 24$, $M = 4.03$, $SD = 0.83$) experienced more Novelty than males ($n = 21$, $M = 3.50$, $SD = 0.87$; $F_{(1,41)} = 5.98$, $p = 0.019$, $\eta_p^2 = 0.13$), regardless of the medium (see Figure 6).

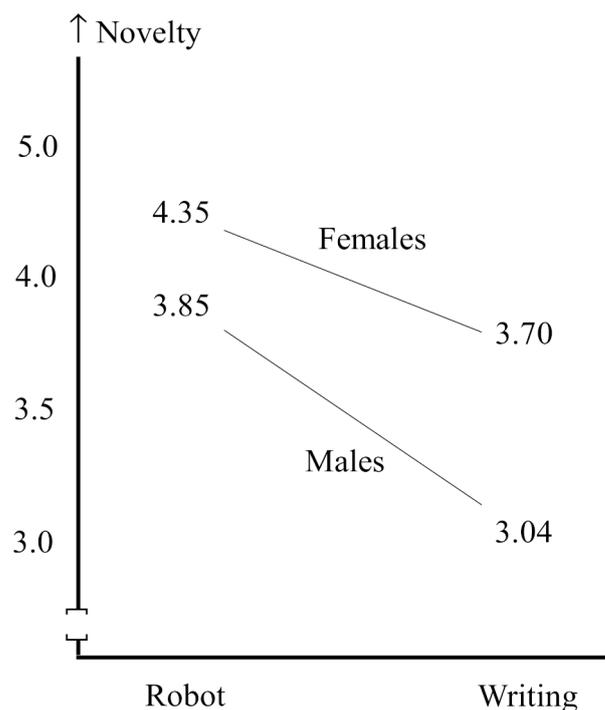


Figure 6. Effects of Gender and Media on Novelty ($N = 45$).

With $n = 31$, only one main effect was significant: Females ($n = 15$, $M = 4.23$, $SD = 0.74$) experienced more Novelty than males ($n = 16$, $M = 3.51$, $SD = 0.95$; $F_{(1,27)} = 5.35$, $p = 0.029$, $\eta_p^2 = 0.17$) and medium did not show significant effects any more ($F_{(1,27)} = 2.98$, $p = 0.95$). Overall, females experienced more novelty, but not particularly with respect to robots.

4. Discussion

Our manipulation was successful: The video was rated as significantly inducing a strong negative mood. Our treatment also was successful: We could demonstrate significant improvement of positive affect and reduction of negative affect after treatment.

We assumed that, after negative-mood induction, (H1) a social robot that invites self-disclosure will lower the level of negative emotions more than writing a journal page. Indeed, our self-disclosure AI chatbot, in unison with the DARwIn Mini embodiment, led to viewers of video recordings of the Wenchuan Sichuan earthquake in China 2008 being significantly more positive. This was particularly so for people who were most negatively affected by the video. For those less affected, writing a diary page also sufficed.

In our study, valence should be conceived of as a bipolar dimension. Significant and reasonably strong main effects of robots exerting more positive results than writing were established by assessing bipolar valence, particularly for participants who responded as experiencing high levels of negativity. Even when we analyzed valence as a within-factor at two levels measured as separate unipolar scales, the significant effects of media

occurred across positive and negative valence, not to these measures separately. Novelty of the medium (either robot or writing) did not affect the effects on bipolar valence or, occasionally, for the less-affected.

Now, a note on the analysis: If we had followed conventional statistical practices, we would have eliminated the outliers from our data set and found no differences between writing and robots in alleviating stress and anxiety. In reporting a null effect, looking at normal distributions only would have missed the upshot that those who were most in need of mental support should not be deprived of a treatment that is more effective than traditional text writing; that is, something that comes closer to a therapist, such as a social robot, which can relieve the shortage of caregivers in the mental care sector.

We also hypothesized (H2) that a social robot that invites self-disclosure is more relevant to goals, needs and concerns than writing on paper; this was not the case. Even though we measured the highest grand mean averages for relevance, whether tested for those high or low in terms of emotional negativity, men or women, the relevance did not differ for any of the tested fixed factors and did not significantly contribute to the effects on valence. It was only in unison with novelty that relevance took effect. The novelty aspect of talking to a robot or writing on paper apparently made the medium more relevant to personal goals and concerns. Furthermore, women experienced more novelty of the presented medium than men; however, this was not specific to the robot or the writing condition.

New technologies, such as social robots, can provide various opportunities for discovering new methods to improve an individual's well-being, suggesting that such new technologies can alleviate the current pressure on healthcare services, such as care for older adults, depressed youth and groups with special needs [83]. Our study focused on social robots helping individuals to improve their mental well-being through self-disclosure. The results suggested that individuals who have a relatively high level of negative emotions benefitted the most from the robot interaction.

Our results were not consistent with [84], where the effect of four conditions of disclosure (writing, private spoken, talking to a passive listener and talking to an active facilitator) had about the same effect in reducing negative emotions as, after the disclosure session, the negative emotions remained. Our results also run counter with [46], where the two procedures (talking and writing) were almost identical in reducing negative affect and in producing adaptive changes in cognition and self-esteem. However, our results were not at odds with [85], who compared writing and talking with a psychotherapist and found that, after writing, no increase of positive emotions happened, whereas, after talking with the therapist, positive emotions increased. Maybe the answer lies in the change of focus: Talking to a (virtual) therapist does not so much decrease negativity, as it compensates negativity by increasing positive affect.

Not reducing negativity, however, would go against the studies by [46,86–88], which all showed that emotional disclosure interventions are effective in reducing the level of negative emotions. Perhaps the decrease in negativity takes longer than the immediate joy of encountering a (virtual) human. The length of the emotional disclosure session in the said studies was much longer than our 10 min: Frattaroli [38] concluded, from a meta-analysis, that such sessions usually last for days or weeks.

Limitations

A limitation of our current study was that participants took the questionnaire only once, after talking with the robot or after writing, rather than after the video as well as after treatment. Emotions are short-lived, notoriously ephemeral and decay rapidly after elicitation [22,89,90]. Self-reported questionnaires should be administered as soon as possible, in order to certify that the measured emotion is related to the experienced one [91,92].

Additionally, emotions tend to subside once individuals reflect on them [93], for example, when filling out a questionnaire. In the future, we may combine self-reported

questionnaires together with physiological reactions and behaviors, for triangulation purposes [9,22,27,94–96].

Our results fall in line with the positive effects exerted by robots in the HRI studies conducted by [51–53,59,60]. In psychotherapy as well, studies have reported on the beneficial effects of robot-enhanced therapy [51,61–64].

Our sample was limited to the Chinese student community, which confines the generalizability of our results, but which did methodologically provide for a homogeneous group. Moreover, the current study was supported by a Hong Kong government project to alleviate the depressed mood which many Hong Kong youth (i.e., students) are experiencing at present and, so, there was also a social reason to study the student community.

5. Conclusions

In conclusion, the situation we were interested in was when no human is available to talk to (e.g., during social isolation due to COVID-19). This was from the understanding that human support topples robot help, no matter what. Commonly, in the absence of human companionship, people may write in a diary or journal, which has been recommended by psychiatrists and psychotherapy studies alike. It is an open-ended, unstructured exercise to reduce stress, not following any protocols. However, what if we could imitate a human psychiatrist with a humanoid robot that follows a certain protocol? Would that help better than writing in a journal? Our study indicates that, for those experiencing strong negative emotions, it does.

The key contributions of our study are as follows: (1) Social robots exert positive therapeutic effects, which are even better than writing for those who are in a very bad mood. (2) Moreover, the robot does not have to be expensive or fancy: The bodywork we used was a simple DARwIn Mini and the chatbot was well-devised, but not expensive. This makes robot companionship accessible to a large audience. (3) As another key element, executing Null-hypothesis testing in a mechanical way may overlook the occurrence of wholesome effects of treatment in a sensitive sub-sample: A cure for cancer may not be effective for those who are not ill, but may save the lives of those who are. Likewise, those who do not feel depressed do not need a robot, while those who are sad and alone are thankful for talking to ‘somebody’.

It will not be hard to apply our system to real-life situations: older adults, depressed youth and forensic patients tend to feel lonesome and may want to self-disclose. Training relevant AI systems can be carried out, using more data. Other robot embodiments can also be employed, and the one main hurdle is the poor speech recognition we have developed so far, which may improve over time.

5.1. Future Work

It may be good to follow up with studies that compare robot to human performance as a ‘benchmark’ for the robot, in terms of what it can and cannot be used for. Furthermore, in comparison to journal writing, it would be interesting to explore which design features were specifically responsible for the effects we found: Was it the robot embodiment or, for example, the type of questions it posed?

User engagement with a robot may be enhanced by a robot’s movements, gaze, non-verbal expressions, better speech and so on. It may be that, through enhanced engagement, the user can feel less depressed. Our focus was on negative-mood reduction (or positive-mood compensation) directly, measuring valence before and after robot interaction. We did not measure the level of engagement and, so, we do not know how engaged the participants were. Including a measure of engagement may help to explain why the robot condition worked better for participants with highly negative emotion. That would be an extra research feature, but does not invalidate the results of the current study, in that negative valence was compensated for by an increase in positive mood.

When evaluating emotive stimuli, it is important to gather different types of data, in order to reify the conclusions drawn from the data, as emotions are subjective. Method-

ologically, as post-experiment questionnaires tend to differ from real-time data analysis, it may be that gathering biometric data can provide a way to sustain our outcomes when employing positive and negative video stimuli. The point, of course, is whether we are interested in immediate negative-mood reduction or wish to investigate more long-term effects. It also may be that wearing certain equipment may interfere with the experience of writing or the robot interaction, which would also be worth investigating. For future research, biometrics pose exciting prospects for triangulation purposes, as well as serving as a topic of methodological investigation in its own right.

5.2. Design Practice

With accumulated evidence for robot-supported mental well-being, we felt that we should set out to put this knowledge into design practice. Currently, we are in the process of developing our own robot, MEME: a talking stress ball that embodies our self-disclosure AI chatbot. We hope that MEME may help people who feel depressed during social isolation, or may calm down those who seek violence to settle their disputes in Hong Kong. Figure 7 shows the development steps so far.

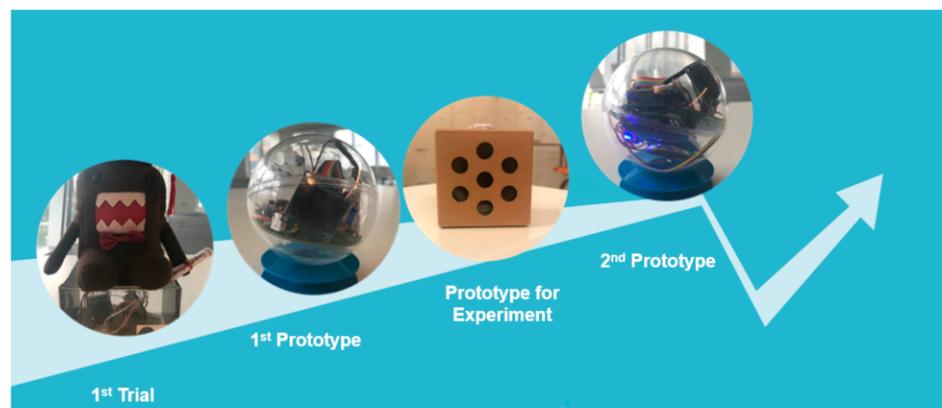


Figure 7. Prototyping the talking stress ball MEME.

MEME is round and covered in silicone rubber for a soft touch, look and feel. It is portable, pocket-size and is easy to carry. The cover can be adapted, according to the user's taste. Interactions with MEME take place through emojis (Figure 8). Many studies in Computer-Mediated Communication have asserted the importance of emojis in non-verbal interactions, in terms of representing a person's affectionate or depressed feelings (see, e.g., [97–100]). MEME can be found at www.roboticmeme.com (accessed on 25 July 2019). The logo was designed according to the directions of [101].

In this study, video footage of the 2008 Sichuan earthquake aroused negative emotions, which were mitigated by self-disclosure to a robot or by writing a journal diary page. The choice of medium was indifferent to most participants: both means worked for them. For those who felt extremely bad after the shocking video, however, the medium did make a difference. Those high on negative valence were significantly more positive after talking to a social robot.

Valence, in this study, was conceived of as a bipolar scale (i.e., more positive is less negative). Relevance had little to do with these effects, and was most susceptible to the novelty of the medium. The newer it is, the more personally relevant the medium seemed. The experience of novelty had little effect on valence, and was higher for robots than for writing. Even though females experienced more novelty throughout, this had nothing to do with robots, as such. Robots seem to be good candidates to aid people with stress and anxiety problems. We took a shot at such opportunities by designing our own MEME stress ball, featuring an emoji-enhanced self-disclosure inviting AI chatbot.

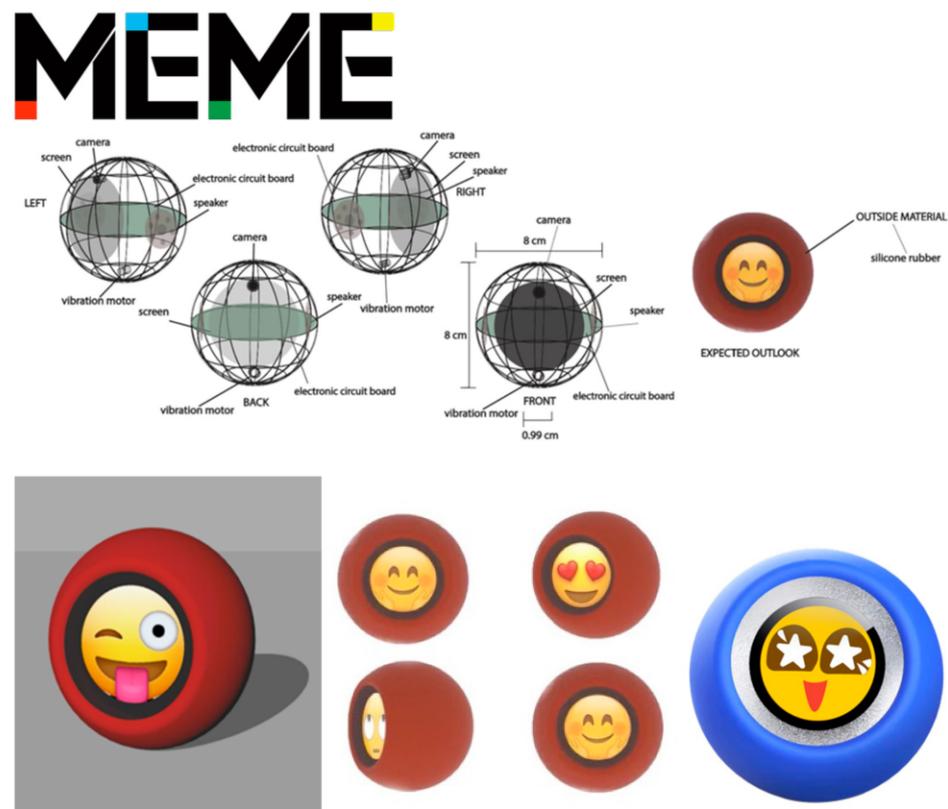


Figure 8. MEME complements the chatbot function with emoticons.

It is noteworthy, however, that the positive effects we observed were valid for the robot as a whole. We should not attribute the positive effects on emotional valence to particular design features, such as specific parts of the embodiment or the quality of the chatbot. As one of the participants astutely commented:

The robot answered my questions in weird ways sometimes, and repeated some questions. I think the unexpected movement of the robot was the best part of the experiment. It affectively changed my mood. Not so much the conversation itself.

Let this be a reminder to us robot researchers, AI developers and designers: The robot made funny moves, which cheered this participant up—not the conversation about difficult things. Perhaps in the future, in concert with talking stress balls, we should create paper and pens that make sudden funny moves, as well.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/robotics10030098/s1>, Technical Report S1: Self-disclosure to a Robot or on Paper.

Author Contributions: Conceptualization, Y.E.D.; methodology, Y.D., M.J.Y. and J.F.H.; software, Z.E.L.; validation, Y.E.D., M.J.Y. and Z.E.L.; formal analysis, Y.E.D., M.J.Y. and J.F.H.; investigation, Y.E.D., M.J.Y. and Z.E.L.; resources, Y.E.D., M.J.Y., Z.E.L. and J.F.H.; data curation, J.F.H.; writing—original draft preparation, Y.D. and J.F.H.; writing—review and editing, J.F.H.; visualization, Y.E.D., M.J.Y. and J.F.H.; supervision, J.F.H.; project administration, J.F.H.; funding acquisition, J.F.H. All authors have read and agreed to the published version of the manuscript.

Funding: The contribution by Johan F. Hoorn was supported by the project Negative-mood reduction among HK youth with robot PAL (Personal Avatar for Life) of the Artificial Intelligence in Design Laboratory in Hong Kong (grant number: AiDLab RP2P3).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Human Subjects Ethics Sub-committee of the university filed under HSEARS20200204003.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data can be made available by the corresponding author.

Acknowledgments: This study was a Chinese, Korean and Dutch collaboration of students in the School of Design and the Dept. of Computing. Zachary Tan is kindly acknowledged for assembling the DARwIn Mini as part of his primary school project on robotics. Jiayuan Chen translated the items of Figure 2 into English. The anonymous reviewers are greatly thanked for helping to improve an earlier draft of our manuscript.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Appendix A

Structured questionnaire for self-disclosure to a robot (English translated from the Chinese). Supplementary Materials S1 provides the questionnaire versions for the robot as well as for writing, in Chinese and English.

Dear Sir/Madam,

Thank you for your time for our experiment. We would like to ask you to answer a few questions. Answering these questions will only take a few minutes.

You have the right to withdraw at any point during the study, for any reason, and without any prejudice. If you would like to contact the Principal Investigator in the study to discuss this research, please e-mail <name> via <name>@connect.polyu.hk.

By clicking the button below, you acknowledge that your participation in the study is voluntary, you are 18 years of age, and that you are aware that you may choose to terminate your participation in the study at any time and for any reason. The data provided by the participants of the study will be processed and published anonymously in the results sections of the paper.

This study is supervised by The Hong Kong Polytechnic University.

Thank you for your participation.

With kind regards,
Team Social Robot MEME

I agree to participate in this study
I do not agree to participate in this study

Scale Valence-before-treatment

I. After seeing the film samples . . .

Vb1i→I feel good

Totally—Disagree—Agree a—Totally
Disagree—Disagree—A Little—Little—Agree—Agree
1—2—3—4—5—6

Vb2i→I am well

Totally—Disagree—Agree a—Totally
Disagree—Disagree—A Little—Little—Agree—Agree
1—2—3—4—5—6

Vb3i→I have positive feelings

Vb4i→I am optimistic

Vb5c→I feel bad

Vb6c→I am unwell

Vb7c→I have negative feelings

Vb8c→I am pessimistic

Scale Valence-after-treatment

II. After talking to the robot . . .

Vb1i→I feel good

Vb2i→I am well

Vb3i→I have positive feelings

Vb4i→I am optimistic

Vb5c→I feel bad

Vb6c→I am unwell

Vb7c→I have negative feelings

Vb8c→I am pessimistic

Scale Relevance

III. To regulate my emotions, talking to the robot is . . .

Re1i→useful

Re2i→worthwhile

Re3c→worthless

Re4c→useless

Scale Novelty

IV. Talking to a robot is . . .

No1i→novel

No2i→original

No3i→unexpected

No4c→predictable

No5c→commonplace

No6c→old-fashioned

Demographics

De1→Gender

Female

Male

Other

De2→Age

De3→What is your highest completed education or current education level?

Primary school or below

Secondary school

Post-secondary school/Associate Degree/Diploma

University undergraduate

Master's degree

Doctoral degree or above

De4→Ethnicity

Asia

Africa

Europe

North America

South America

Australia/Oceania

Antarctica

If you have any further questions or remarks about this questionnaire,

please let us know.
You can write your feedback below.

Kind regards,

Social Robot MEME

<name>@connect.polyu.hk

References

- Bu, F.; Steptoe, A.; Fancourt, D. Who is lonely in lockdown? Cross-cohort analyses of predictors of loneliness before and during the COVID-19 pandemic. *Public Health* **2020**, *186*, 31–34. [[CrossRef](#)]
- Bu, F.; Steptoe, A.; Fancourt, D. Loneliness during a strict lockdown: Trajectories and predictors during the COVID-19 pandemic in 38,217 United Kingdom adults. *Soc. Sci. Med.* **2020**, *265*, 113521. [[CrossRef](#)]
- Killgore, W.D.; Cloonan, S.A.; Taylor, E.C.; Miller, M.A.; Dailey, N.S. Three months of loneliness during the COVID-19 lockdown. *Psychiatry Res.* **2020**, *293*, 113392. [[CrossRef](#)]
- Li, L.Z.; Wang, S. Prevalence and predictors of general psychiatric disorders and loneliness during COVID-19 in the United Kingdom. *Psychiatry Res.* **2020**, *291*, 113267. [[CrossRef](#)]
- Tso, I.F.; Park, S. Alarming levels of psychiatric symptoms and the role of loneliness during the COVID-19 epidemic: A case study of Hong Kong. *Psychiatry Res.* **2020**, *293*, 113423. [[CrossRef](#)] [[PubMed](#)]
- Williams, C.Y.; Townson, A.T.; Kapur, M.; Ferreira, A.F.; Nunn, R.; Galante, J.; Phillips, V.; Gentry, S.; Usher-Smith, J.A. Interventions to reduce social isolation and loneliness during COVID-19 physical distancing measures: A rapid systematic review. *PLoS ONE* **2021**, *16*, e0247139. [[CrossRef](#)] [[PubMed](#)]
- Hopwood, T.L.; Schutte, N.S. Psychological outcomes in reaction to media exposure to disasters and large-scale violence: A meta-analysis. *Psychol. Violence* **2017**, *7*, 316–327. [[CrossRef](#)]
- Twenge, J.M.; Joiner, T.E.; Rogers, M.L.; Martin, G.N. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. *Clin. Psychol. Sci.* **2018**, *6*, 3–17. [[CrossRef](#)]
- Frijda, N. *The Laws of Emotion*; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2007.
- Lerner, J.; Li, Y.; Valdesolo, P.; Kassam, K. Emotion and decision making. *Annu. Rev. Psychol.* **2015**, *66*, 799–823. [[CrossRef](#)]
- Croyle, R.T.; Uretzky, M.D. Effects of mood on self-appraisal of health status. *Health Psychol.* **1987**, *6*, 239–253. [[CrossRef](#)]
- Levine, L.; Edelman, R. Emotion and memory narrowing: A review and goal relevance approach. *Cogn. Emot.* **2009**, *23*, 833–875. [[CrossRef](#)]
- Mayne, T.J. Emotions and health. In *Emotions: Current Issues and Future Directions*; Mayne, T.J., Bonanno, G.A., Eds.; Guilford: New York, NY, USA, 2001; pp. 361–397.
- Thayer, J.F.; Ruiz-Padial, E. Neurovisceral integration, emotions and health: An update. *Int. Congr. Ser.* **2006**, *1287*, 122–127. [[CrossRef](#)]
- Pressman, S.; Gallagher, M.; Lopez, S. Is the emotion-health connection a “first-world problem”? *Psychol. Sci.* **2013**, *24*, 544–549. [[CrossRef](#)] [[PubMed](#)]
- Consedine, N.S.; Moskowitz, J.T. The role of discrete emotions in health outcomes: A critical review. *Appl. Prev. Psychol.* **2007**, *12*, 59–75. [[CrossRef](#)]
- Mayne, T.J. Negative affect and health: The importance of being earnest. *Cogn. Emot.* **1999**, *13*, 601–635. [[CrossRef](#)]
- Kubzansky, L.D.; Kawachi, I. Going to the heart of the matter: Do negative emotions cause coronary heart disease? *J. Psychosom. Res.* **2000**, *48*, 323–337. [[CrossRef](#)]
- Sonnemans, J.; Frijda, N. The determinants of subjective emotional intensity. *Cogn. Emot.* **1995**, *9*, 483–506. [[CrossRef](#)]
- Lazarus, R. *Emotion and Adaptation*; Oxford University: Oxford, UK, 1991.
- Parkinson, B. *Ideas and Realities of Emotion*; Routledge: London, UK, 1995.
- Scherer, K.R. What are emotions? And how can they be measured? *Soc. Sci. Inf.* **2005**, *44*, 695–729. [[CrossRef](#)]
- Russell, J.A. Core affect and the psychological construction of emotion. *Psychol. Rev.* **2003**, *110*, 145–172. [[CrossRef](#)]
- Shuman, V.; Sander, D.; Scherer, K.R. Levels of valence. *Front. Psychol.* **2013**, *4*, 261. [[CrossRef](#)]
- Barrett, L. Valence is a basic building block of emotional life. *J. Res. Pers.* **2006**, *40*, 35–55. [[CrossRef](#)]
- Scherer, K.R. Appraisal considered as a process of multilevel sequential checking. In *Appraisal Processes in Emotion: Theory, Methods, Research*; Scherer, K.R., Schorr, A., Johnstone, T., Eds.; Oxford University: New York, NY, USA, 2001; pp. 92–120.
- Frijda, N. *The Emotions (Studies in Emotion and Social Interaction)*; Cambridge University: New York, NY, USA, 1986.
- Scherer, K.R. The nature and dynamics of relevance and valence appraisals: Theoretical advances and recent evidence. *Emot. Rev.* **2013**, *5*, 150–162. [[CrossRef](#)]
- Scherer, K.R. On the nature and function of emotions: A component process approach. In *Approaches to Emotion*; Scherer, K.R., Ekman, P., Eds.; L. Erlbaum Associates: Hillsdale, NJ, USA, 1984; pp. 293–317.
- Fontaine, J.; Scherer, K.; Soriano, C. *Components of Emotional Meaning. A Sourcebook*; Oxford University: Oxford, UK, 2013.
- Farber, B. *Self-Disclosure in Psychotherapy*; Guilford: New York, NY, USA, 2006.

32. Smyth, J. Written emotional expression: Effect sizes, outcome types, and moderating variables. *J. Consult. Clin. Psychol.* **1998**, *66*, 174–184. [CrossRef] [PubMed]
33. Pennebaker, J.W. Traumatic experience and psychosomatic disease. Exploring the roles of behavioural inhibition, obsession, and confiding. *Can. Psychol.* **1985**, *26*, 82–95. [CrossRef]
34. Alford, W.; Malouff, J.; Osland, K. Written emotional expression as a coping method in child protective services officers. *Int. J. Stress Manag.* **2005**, *12*, 177–187. [CrossRef]
35. Hemenover, S. Individual differences in rate of affect change: Studies in affective chronometry. *J. Pers. Soc. Psychol.* **2003**, *85*, 121–131. [CrossRef] [PubMed]
36. Horneffer, K.; Jamison, P. The emotional effects of writing about stressful experiences: An exploration of moderators. *Occup. Health Care* **2002**, *16*, 77–89. [CrossRef]
37. Ireland, M.; Malouff, J.; Byrne, B. The efficacy of written emotional expression in the reduction of psychological distress in police officers. *Int. J. Police Sci. Manag.* **2007**, *9*, 303–311. [CrossRef]
38. Frattaroli, J. Experimental disclosure and its moderators: A meta-analysis. *Psychol. Bull.* **2006**, *132*, 823–865. [CrossRef] [PubMed]
39. Pennebaker, J.; Francis, M. Cognitive, emotional, and language processes in disclosure. *Cogn. Emot.* **1996**, *10*, 601–626. [CrossRef]
40. Dalton, J.; Glenwick, D. Effects of expressive writing on standardized graduate entrance exam performance and physical health functioning. *J. Psychol.* **2009**, *143*, 279–292. [CrossRef] [PubMed]
41. Frisina, P.G.; Borod, J.C.; Lepore, S.J. A meta-analysis of the effects of written emotional disclosure on the health outcomes of clinical populations. *J. Nerv. Ment. Dis.* **2004**, *192*, 629–634. [CrossRef] [PubMed]
42. Pascoe, P. Using patient writings in psychotherapy: Review of evidence for expressive writing and cognitive-behavioral writing therapy. *Am. J. Psychiatry Resid. J.* **2016**, *11*, 3–6. [CrossRef]
43. Lumley, M. Alexithymia, emotional disclosure, and health: A program of research. *J. Pers.* **2004**, *72*, 1271–1300. [CrossRef] [PubMed]
44. Pennebaker, J. Putting stress into words: Health, linguistic, and therapeutic implications. *Behav. Res.* **1993**, *31*, 539–548. [CrossRef]
45. Pennebaker, J.; Beall, S. Confronting a traumatic event: Toward an understanding of inhibition and disease. *J. Abnorm. Psychol.* **1986**, *95*, 274–281. [CrossRef]
46. Murray, E.; Segal, D. Emotional processing in vocal and written expression of feelings about traumatic experiences. *J. Trauma. Stress* **1994**, *7*, 391–405. [CrossRef]
47. World Health Organization. WHO's Mental Health Atlas 2017 Highlights Global Shortage of Health Workers Trained in Mental Health. Available online: <https://www.who.int/hrh/news/2018/WHO-MentalHealthAtlas2017-highlights-HW-shortage/en/> (accessed on 6 June 2018).
48. Weizenbaum, J. ELIZA—A computerS program for the study of natural language communication between man and machine. *Commun. ACM* **1966**, *9*, 36–45. [CrossRef]
49. Saffarizadeh, K.; Boodraj, M.; Alashoor, T.M. Conversational assistants: Investigating privacy concerns, trust, and self-disclosure. In Proceedings of the International Conference on Information System: Transforming Society with Digital Innovation (ICIS '17), Seoul, Korea, 10–13 December 2017; Kim, Y.J., Agarwal, R., Lee, J.K., Eds.; Association for Information Systems: Atlanta, GA, USA, 2017.
50. Hoorn, J.F.; Konijn, E.A.; Germans, D.M.; Burger, S.; Munneke, A. The in-between machine: The unique value proposition of a robot or why we are modelling the wrong things. In Proceedings of the 7th International Conference on Agents and Artificial Intelligence (ICAART), Lisbon, Portugal, 10–12 January 2015; Loiseau, S., Filipe, J., Duval, B., van den Herik, J., Eds.; ScitePress: Lisbon, Portugal, 2015; pp. 464–469.
51. Wada, K.; Shibata, T.; Saito, T.; Sakamoto, K.; Tanie, K. Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA '05), Barcelona, Spain, 18–22 April 2005; pp. 2785–2790. [CrossRef]
52. Jibb, L.A.; Birnie, K.A.; Nathan, P.C.; Beran, T.N.; Hum, V.; Victor, J.C.; Stinson, J.N. Using the MEDiPORT humanoid robot to reduce procedural pain and distress in children with cancer: A pilot randomized controlled trial. *Pediatr. Blood Cancer* **2018**, *65*, e27242. [CrossRef] [PubMed]
53. Dang, T.-H.-H.; Tapus, A. The role of motivational robotic assistance in reducing user's task stress. *Int. J. Soc. Robot.* **2015**, *7*, 227–240. [CrossRef]
54. Cabibihan, J.; Javed, H.; Ang, M.; Aljunied, S. Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *Int. J. Soc. Robot.* **2013**, *5*, 593–618. [CrossRef]
55. Robins, B.; Amirabdollahian, F.; Ji, Z.; Dautenhahn, K. Tactile interaction with a humanoid robot for children with autism: A case study analysis involving user requirements and results of an initial implementation. In Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '10), Viareggio, Italy, 13–15 September 2010; Avizzano, C.A., Ruffaldi, E., Eds.; IEEE: Piscataway, NJ, USA, 2010; pp. 704–711. [CrossRef]
56. Kozima, H.; Nakagawa, C.; Yasuda, Y. Interactive robots for communication care: A case-study in autism therapy. In Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication (ROMAN '05), Nashville, TN, USA, 13–15 August 2005; IEEE: Piscataway, NJ, USA, 2005; pp. 341–346.
57. Vanderborght, B.; Simut, R.; Saldien, J.; Pop, C.; Rusu, A.S.; Pintea, S.; Dirk, L.; David, D.O. Using the social robot Probo as a social story telling agent for children with ASD. *Interact. Stud. Soc. Behav. Commun. Biol. Artif. Syst.* **2012**, *13*, 348–372. [CrossRef]

58. Tapus, A.; Peca, A.; Aly, A.; Pop, C.; Jisa, L.; Pinte, S.; David, D.O. Children with autism social engagement in interaction with Nao, an imitative robot. A series of single case experiments. *Interact. Stud. Soc. Behav. Commun. Biol. Artif. Syst.* **2012**, *13*, 315–347. [[CrossRef](#)]
59. Kumazaki, H.; Warren, Z.; Swanson, A.; Yoshikawa, Y.; Matsumoto, Y.; Takahashi, H.; Kikuchi, M. Can robotic systems promote self-disclosure in adolescents with autism spectrum disorder? A pilot study. *Front. Psychiatry* **2018**, *9*, 36. [[CrossRef](#)]
60. Pu, L.; Moyle, W.; Jones, C.; Todorovic, M. The effectiveness of social robots for older adults: A systematic review and meta-analysis of randomized controlled studies. *Gerontologist* **2019**, *59*, e37–e51. [[CrossRef](#)]
61. Libin, A.V.; Libin, E.V. Person-robot interactions from the robopsychologists' point of view: The robotic psychology and robototherapy approach. *Proc. IEEE* **2004**, *92*, 1789–1803. [[CrossRef](#)]
62. Costescu, C.A.; Vanderborght, B.; David, D.O. The effects of robot-enhanced psychotherapy: A meta-analysis. *Rev. Gen. Psychol.* **2014**, *18*, 127–136. [[CrossRef](#)]
63. Kidd, C.; Taggart, W.; Turkle, S. A sociable robot to encourage social interaction among the elderly. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '06), Orlando, FL, USA, 15–19 May 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 3972–3976.
64. Wada, K.; Shibata, T. Social effects of robot therapy in a care house-change of social network of the residents for two months. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation (ICRA '07), Roma, Italy, 10–14 April 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 1250–1255.
65. Bolls, P.; Lang, A.; Potter, R. The effects of message valence and listener arousal on attention, memory, and facial muscular responses to radio advertisements. *Commun. Res.* **2001**, *28*, 627–651. [[CrossRef](#)]
66. Lang, A.; Shin, M.; Lee, S. Sensation seeking, motivation, and substance use: A dual system approach. *Media Psychol.* **2005**, *7*, 1–29. [[CrossRef](#)]
67. Siedlecka, E.; Denson, T. Experimental methods for inducing basic emotions: A qualitative review. *Emot. Rev.* **2019**, *11*, 87–97. [[CrossRef](#)]
68. Martelaro, N.; Nneji, V.; Ju, W.; Hinds, P. Tell me more designing HRI to encourage more trust, disclosure, and companionship. In Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI '16), Christchurch, New Zealand, 7–10 March 2016; Bartneck, C., Nagai, Y., Paiva, A., Šabanović, S., Eds.; IEEE: Piscataway, NJ, USA, 2016; pp. 181–188.
69. Złotowski, J.; Sumioka, H.; Nishio, S.; Glas, D.F.; Bartneck, C.; Ishiguro, H. Appearance of a robot affects the impact of its behaviour on perceived trustworthiness and empathy. *Paladyn. J. Behav. Robot.* **2016**, *7*, 55–66. [[CrossRef](#)]
70. Salem, M.; Lakatos, G.; Amirabdollahian, F.; Dautenhahn, K. Would you trust a (faulty) robot: Effects of error, task type and personality on human-robot cooperation and trust. In Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15), Portland, OR, USA, 2–5 March 2015; ACM: New York, NY USA, 2015; pp. 141–148. [[CrossRef](#)]
71. Nystul, M. *Introduction to Counseling: An Art and Science Perspective*, 5th ed.; Sage: Thousand Oaks, CA, USA, 2016.
72. Altman, I.; Taylor, D. *Social Penetration: The Development of Interpersonal Relationships*; Holt: New York, NY, USA, 1973.
73. Nomura, T.; Kawakami, K. Relationships between robot's self-disclosures and human's anxiety toward robots. In Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI-IAT'11), Lyon, France, 22–27 August 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 66–69. [[CrossRef](#)]
74. Psychopathology Committee of the Group for the Advancement of Psychiatry Reexamination of therapist self-disclosure. *Psychiatr. Serv.* **2001**, *52*, 1489–1493. [[CrossRef](#)]
75. Smith, C.A.; Ellsworth, P.C. Patterns of cognitive appraisal in emotions. *J. Personal. Soc. Psychol.* **1985**, *48*, 813–838. [[CrossRef](#)]
76. Carrera, P.; Oceja, L. Drawing mixed emotions: Sequential or simultaneous experiences? *Cogn. Emot.* **2007**, *21*, 422–441. [[CrossRef](#)]
77. Russell, J.A. Mixed emotions viewed from the psychological constructionist perspective. *Emot. Rev.* **2017**, *9*, 111–117. [[CrossRef](#)]
78. Watson, D.; Clark, L.A.; Tellegen, A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Personal. Soc. Psychol.* **1988**, *54*, 1063–1070. [[CrossRef](#)]
79. Diener, E. Introduction to the special section on the structure of emotion. *J. Personal. Soc. Psychol.* **1999**, *76*, 803–804. [[CrossRef](#)]
80. Russell, J.A.; Carroll, J.M. On the bipolarity of positive and negative affect. *Psychol. Bull.* **1999**, *125*, 3–30. [[CrossRef](#)]
81. Russell, J.A.; Carroll, J.M. The phoenix of bipolarity: Reply to Watson and Tellegen. *Psychol. Bull.* **1999**, *125*, 611–617.
82. Friedman, L.M.; Furberg, C.D.; DeMets, D.L. *Fundamentals of Clinical Trials*; Springer: New York, NY, USA, 2010.
83. Broadbent, E. Interactions with robots: The truths we reveal about ourselves. *Annu. Rev. Psychol.* **2017**, *68*, 627–652. [[CrossRef](#)] [[PubMed](#)]
84. Slavin-Spenny, O.; Cohen, J.; Oberleitner, L.; Lumley, M. The effects of different methods of emotional disclosure: Differentiating post-traumatic growth from stress symptoms. *J. Clin. Psychol.* **2011**, *67*, 993–1007. [[CrossRef](#)]
85. Murray, E.J.; Lamnin, A.D.; Carver, C.S. Emotional expression in written essays and psychotherapy. *J. Soc. Clin. Psychol.* **1989**, *8*, 414–429. [[CrossRef](#)]
86. Epstein, E.M.; Sloan, D.P.; Marx, B. Getting to the heart of the matter: Written disclosure, gender, and heart rate. *Psychosom. Med.* **2005**, *67*, 413–419. [[CrossRef](#)]
87. Sloan, D.M.; Marx, B.P.; Epstein, E.M.; Lexington, J.M. Does altering the writing instructions influence outcome associated with written disclosure? *Behav. Ther.* **2007**, *38*, 155–168. [[CrossRef](#)]

88. Perez, S.; Penate, W.; Bethencourt, J.; Fumero, A. Verbal emotional disclosure of traumatic experiences in adolescents: The role of social risk factors. *Front. Psychol.* **2017**, *8*, 372. [[CrossRef](#)]
89. Costa, T.; Cauda, F.; Crini, M.; Tatu, M.; Celeghin, A.; De Gelder, B.; Tamietto, M. Temporal and spatial neural dynamics in the perception of basic emotions from complex scenes. *Soc. Cogn. Affect. Neurosci.* **2014**, *9*, 1690–1703. [[CrossRef](#)]
90. Hassenzahl, M. Emotions can be quite ephemeral; we cannot design them. *Interactions* **2004**, *11*, 46–48. [[CrossRef](#)]
91. Lane, A.M.; Terry, P.C. The nature of mood: Development of a conceptual model with a focus on depression. *J. Appl. Sport Psychol.* **2000**, *12*, 16–33. [[CrossRef](#)]
92. Qiao-Tasserit, E.; Garcia Quesada, M.; Antico, L.; Bavelier, D.; Vuilleumier, P.; Pichon, S. Transient emotional events and individual affective traits affect emotion recognition in a perceptual decision-making task. *PLoS ONE* **2017**, *12*, e0171375. [[CrossRef](#)]
93. Mauss, I.; Robinson, M. Measures of emotion: A review. *Cogn. Emot.* **2009**, *23*, 209–237. [[CrossRef](#)] [[PubMed](#)]
94. Erevelles, S. The role of affect in marketing. *J. Bus. Res.* **1998**, *42*, 199–215. [[CrossRef](#)]
95. Lang, P. The three-system approach to emotion. In *The Structure of Emotion*; Birbaum, N., Ohman, A., Eds.; Hogrefe: Bern, Switzerland, 1993; pp. 18–30.
96. Scherer, K.R.; Zentner, M.R. Emotional effects of music: Production rules. In *Music and Emotion: Theory and Research*; Juslin, P.N., Sloboda, J.A., Eds.; Oxford University: Oxford, UK, 2001; pp. 361–392.
97. Crystal, D. *Language and the Internet*; Cambridge University: Cambridge, UK, 2006.
98. Rezabek, L.L.; Cochenour, J.J. Visual cues in Computer-Mediated Communication: Supplementing text with emoticons. *J. Vis. Lit.* **1998**, *18*, 201–215. [[CrossRef](#)]
99. Wolf, A. Emotional expression online: Gender difference in emotion use. *CyberPsychol. Behav.* **2000**, *3*, 827–833. [[CrossRef](#)]
100. Li, L.; Yang, Y. Pragmatic functions of emoji in internet-based communication—A corpus-based study. *Asian-Pac. J. Second Foreign Lang. Educ.* **2018**, *3*, 16. [[CrossRef](#)]
101. Walsh, M.F.; Winterich, K.P.; Mittal, V. Do logo redesigns help or hurt your brand? The role of brand commitment. *J. Prod. Brand Manag.* **2010**, *19*, 76–84. [[CrossRef](#)]