Effectiveness of an Empathic Chatbot in Combating Adverse Effects of Social Exclusion on Mood

Mauro de Gennaro[1], Eva G. Krumhuber[1]* and Gale Lucas[2]

[1] Department of Experimental Psychology, University College London, London, United Kingdom, [2] Institute for Creative Technologies, University of Southern California, Los Angeles, CA, United States

From past research it is well known that social exclusion has detrimental consequences for mental health. To deal with these adverse effects, socially excluded individuals frequently turn to other humans for emotional support. While chatbots can elicit social and emotional responses on the part of the human interlocutor, their effectiveness in the context of social exclusion has not been investigated. In the present study, we examined whether an empathic chatbot can serve as a buffer against the adverse effects of social ostracism. After experiencing exclusion on social media, participants were randomly assigned to either talk with an empathetic chatbot about it (e.g., "I'm sorry that this happened to you") or a control condition where their responses were merely acknowledged (e.g., "Thank you for your feedback"). Replicating previous research, results revealed that experiences of social exclusion dampened the mood of participants. Interacting with an empathetic chatbot, however, appeared to have a mitigating impact. In particular, participants in the chatbot intervention condition reported higher mood than those in the control condition. Theoretical, methodological, and practical implications, as well as directions for future research are discussed.

Keywords: social exclusion, empathy, mood, chatbot, virtual human

INTRODUCTION

Experiences of social exclusion, such as being isolated or excluded from a group but not necessarily ignored or explicitly disliked (Williams, 2001; Williams et al., 2005), threaten one of our most basic human needs: the desire to form strong and stable interpersonal relationships with others (Baumeister and Leary, 1995; Gonsalkorale and Williams, 2007). Being on the receiving end of social exclusion and isolation is detrimental as it can lead to a litany of negative consequences ranging from depression (Nolan et al., 2003) and low self-esteem (Leary et al., 1995) to anxiety and perceived lack of control and meaninglessness (Baumeister and Tice, 1990; Zadro et al., 2004; Stillman et al., 2009).

Unfortunately, most individuals at some point in their lives will be both sources and targets of some form of social exclusion (Williams et al., 2005). Indeed, 67% of surveyed Americans admitted giving the silent treatment to a loved one, while 75% reported having received the silent treatment from a loved one (Faulkner et al., 1997). Loneliness or feeling alone (Peplau and Perlman, 1982) are also serious problems plaguing up to a quarter of all Americans multiple times a week (Davis and Smith, 1998), with experiences of social exclusion leading to a downward spiral of further social isolation (see Lucas et al., 2010).

To date, only a few interventions have been designed to help excluded individuals recover from the adverse effects of social exclusion. Researchers have found that some interventions like

emotional support animals (Aydin et al., 2012) *can* help ameliorate the negative impacts of being socially excluded. When people are excluded via social media platforms, new options open up to intervene using technology. As such, it has been shown that an online instant messaging conversation with a stranger improves self-esteem and mood after social exclusion compared to when playing a solitary computer game (Gross, 2009). In this paper, we consider a similar intervention that is more accessible: an empathetic chatbot. Specifically, we tested the possibility that an empathetic chatbot could be used to mitigate the negative impacts of exclusion.

Agents for Mental Health

Virtual humans or agents (i.e., animated characters that allow people to interact with them in a natural way via speech or, in the case of chatbots, via text) have been designed to address various aspects of mental health and social functioning. For example, virtual agents serve as role players in mental health-related applications such as virtual reality exposure therapy (VRET). For this, therapists use agents as part of a scenario designed to evoke clinical symptoms (e.g., PTSD, social anxiety, fear of public speaking) and then guide patients in managing their emotional responses (Jarrell et al., 2006; Anderson et al., 2013; Batrinca et al., 2013). Virtual agents have also been developed to prevent such disorders and symptoms in the first place (e.g., Rizzo et al., 2012). Likewise, virtual agents can help identify disorders and symptoms by interviewing patients about their mental health (e.g., DeVault et al., 2014), and initial evidence suggests that they may be able to evoke greater openness about mental health than human clinical interviewers (Slack and Van Cura, 1968; Lucas et al., 2014; Pickard et al., 2016) or anonymous online clinical surveys (Lucas et al., 2017).

As pointed out by recent research, there is an enormous potential that chatbots hold for addressing mental health-related issues (Følstad and Brandtzaeg, 2017; Brandtzaeg and Følstad, 2018). This bourgeoning field can trace its origins back to the chatbot ELIZA (Weizenbaum, 1967) which imitated Rogerian therapy (Smestad, 2018) by rephrasing many of the statements made by the patient (e.g., if a user were to write "*I have a brother,*" ELIZA would reply "*Tell me more about your brother*"). Following ELIZA, a litany of chatbots and other applications were developed to provide self-guided mental health support for symptom relief (Tantam, 2006). A meta—analysis of 23 randomized controlled trials found that some of these self-guided applications were as effective as standard face-to-face care (Cuijpers et al., 2009). Likewise, embodied conversational agents (ECAs) can be used in cognitive-based therapy (CBT) for addressing anxiety, mood and substance use disorders (Provoost et al., 2017). Chatbots which receive text inputs from users are also beneficial as "virtual agents" in supporting self-guided mental health interventions. For example, the chatbot Woebot (Lim, 2017) guides users through CBT, helping users to significantly reduce anxiety and depression (Fitzpatrick et al., 2017).

These applications can offer help when face-to-face treatment is unavailable (Miner et al., 2016). Additionally, they may assist in overcoming the stigma around mental illness. People expect therapeutic conversational agents to be good listeners,

keep secrets and honor confidentiality (Kim et al., 2018). Since chatbots do not think and cannot form their own judgments, individuals feel more comfortable confiding in them without fear of being judged (Lucas et al., 2014). These beliefs help encourage people to utilize chatbots. As such, participants commonly cite the agents' ability to talk about embarrassing topics and listen without being judgmental (Zamora, 2017).

Agents for Emotional and Social Support

Similar but distinct efforts have been made to develop virtual agents and chatbots for companionship. Much of this work has focused on robotic companions for the elderly. For example, research suggests that Paro – a furry robotic toy seal- may have therapeutic effects which are comparable to live animal therapy. The robot provides companionship to the user by vividly reacting to the user's touch using voices and gestures. Randomized control trials found that Paro reduced stress and anxiety (Petersen et al., 2017) as well as increased social interaction (Wada and Shibata, 2007) in the elderly. Virtual agents, including chatbots, also exist for companionship in older adults (Vardoulakis et al., 2012; Wanner et al., 2017), such as during hospital stays (Bickmore et al., 2015). Moreover, users sometimes form social bonds with agents (i.e., designed for fitness and health purposes) not originally intended for companionship (Bickmore et al., 2005).

Besides their potential for companionship, conversational agents have been developed to provide emotional support. When people experience negative emotions or stress, they often talk to others about their problems and seek comfort from them. Multiple studies have shown that access to support networks has significant health benefits in humans (Reblin and Uchino, 2008). For example, socio-emotional support leads to lower blood pressure (Gerin et al., 1992), reduces the chances of having a myocardial infarction (Ali et al., 2006), decreases mortality rates (Zhang et al., 2007), and helps cancer patients feel more empowered and confident (Ussher et al., 2006).

With regard to human-computer interaction, there is evidence suggesting that chatbots and virtual agents have the potential to reduce negative emotions such as stress (Prendinger et al., 2005; Huang et al., 2015), emotional distress (Klein et al., 2002), and frustration (Hone, 2006), as well as comfort users. Talking to chatbots about negative emotions or stressful experiences may also have benefits over discussing these issues with other humans. Such disclosure to chatbots can have similar emotional, relational and psychological effects as disclosing to another human (Ho et al., 2018) even though – or perhaps because – machines cannot experience emotions or judgment (Bickmore, 2003; Bickmore and Picard, 2004; Lucas et al., 2014; but see also Liu and Sundar, 2018).

Despite people's doubt that machines can have emotional experiences, they typically respond better to agents that express emotions than those that do not (de Melo et al., 2014; Zumstein and Hundertmark, 2017). This fact can be leveraged to make emotional support agents more effective by having them display empathetic responses. In order to comfort someone in a state of grief or distress, it is known that humans employ different communicative behaviors aimed at reducing the emotional distress of another individual (Burleson and Goldsmith, 1996):

one such action is empathic behavior. In the context of human-computer interaction, the empathizing agent communicates his or her understanding of the other individual's emotional state (Reynolds, 2017). Recipients find such responses comforting, resulting in positive impacts on their well-being and health outcomes (Bickmore and Fernando, 2009).

Chatbots may then be able to use empathetic responses to support users just like humans do (Bickmore and Picard, 2005). For example, Brave et al. (2005) found that virtual agents that used empathetic responses were rated as more likeable, trustworthy, caring, and supporting compared to agents that did not employ such responses. As such, the more empathic feedback an agent provides, the more effective it is at comforting users (Bickmore and Schulman, 2007; see also Nguyen and Masthoff, 2009).

Social Exclusion

Interestingly, little to no work has been done to explore the possibility that virtual agents can provide effective support – *specifically* – after experiences of social exclusion. This is surprising given that social exclusion poses a particular risk because, like mental health symptoms, it can culminate over time in pervasive negative consequences. There is a large body of evidence indicating that experiences of social exclusion are associated with detrimental effects such as depression (Nolan et al., 2003), low self-esteem (Leary et al., 1995), feelings of loneliness, helplessness, frustration and jealousy (Leary, 1990; Williams, 2007), anxiety (Baumeister and Tice, 1990), lower control and belonging (Zadro et al., 2004), and reduced perceptions of life as being meaningful (Stillman et al., 2009), Moreover, being repeatedly rejected by others has been associated with attempted suicide (Williams et al., 2005) – and even with school shootings (Leary et al., 2003).

From an evolutionary perspective, it has been argued that humans developed an enhanced sensitivity to detect cues indicative of social exclusion because such state was often a death sentence to our ancestors (Pickett and Gardner, 2005). In fact, group membership enhances the species' survival in non-human primates and animals (Silk et al., 2003), while being ostracized is associated with increased mortality (Lancaster, 1986). It is not surprising then that experiences of social exclusion are painful and distressing, leading to intense negative psychological reactions (Williams, 2007, 2009).

An extensive literature demonstrates that social exclusion causes pain and dampens mood. First, social exclusion is a painful experience which activates similar regions in the cortex as physical pain. For instance, Eisenberger et al. (2003) and Eisenberger and Lieberman (2004) collected fMRI data following experiences of social exclusion and found heightened activation of the dorsal anterior cingulate cortex (dACC), which is also activated during physical pain. Additionally, measurements from ERP, EMG, and EEG confirmed that exclusion has well developed neurobiological foundations (Kawamoto et al., 2013), and through these neurological mechanisms, social exclusion can even cause people to *feel cold* (Zhong and Leonardelli, 2008; IJzerman et al., 2012). Furthermore, social exclusion not only hurts when it comes from loved ones or in-group members,

it is also distressing and painful when the person is excluded by out-group members (Smith and Williams, 2004). More importantly for the current work, research has shown that social exclusion regardless of the source can negatively affect mood (e.g., Gerber and Wheeler, 2009; Lustenberger and Jagacinski, 2010) by lowering positive affect and increasing negative affect (see Williams, 2007 for a review).

Until now, only a handful of studies have explored approaches to help excluded individuals recover from the adverse effects of social exclusion. While some interventions (e.g., emotional support animals or online instant messaging with a stranger, Gross, 2009; Aydin et al., 2012) may help improve self-esteem and mood, no controlled studies yet exist that investigate whether virtual agents could serve as a buffer against the detrimental effects of social exclusion. In this paper, we begin to explore this possibility by testing the effectiveness of an agent (i.e., an empathetic chatbot) in helping people recover from a particular detrimental effect of social exclusion: dampened mood.

Present Research

We begin the foray into this area of research with an empathetic chatbot designed to restore mood after social exclusion. Virtual agents may have the potential to help address this problem. As described above, when chatbots act in the role of humans, they can effectively provide emotional support. This possibility relies on our basic willingness to treat agents like humans. In Media Equation Theory, Nass and colleagues posit that people will respond fundamentally to media (e.g., fictional characters, cartoon depictions, virtual humans) as they would to humans (e.g., Reeves and Nass, 1996; Nass et al., 1997; Nass and Moon, 2000; see also Waytz et al., 2010). For instance, when interacting with an advice-giving agent, users try to be as polite (Reeves and Nass, 1996) as they would with humans. However, such considerations are not afforded to other virtual objects that do not act or appear human (Brave and Nass, 2007; Epley et al., 2007; see also Nguyen and Masthoff, 2009; Khashe et al., 2017). For example, people are more likely to cooperate with a conversational agent that has a human-like face rather than, for instance, an animal face (Friedman, 1997; Parise et al., 1999). Furthermore, it has been found that chatbots with more humanlike appearance make conversations feel more natural (Sproull et al., 1996), facilitate building rapport and social connection (Smestad, 2018), as well as increase perceptions of trustworthiness, familiarity, and intelligence (Terada et al., 2015; Meyer et al., 2016) besides being rated more positively (Baylor, 2011).

Importantly for this work, there is also some suggestion that virtual agents might be capable of addressing a person's need to belong like humans do (Krämer et al., 2012). For example, Krämer et al. (2018) demonstrated that people feel socially satiated after interacting with a virtual agent, akin to when reading a message from a loved one (Gardner et al., 2005). Because they are "real" enough to many of us psychologically, empathetic virtual agents and chatbots may often provide emotional support with greater psychological safety (Kahn, 1990). Indeed, when people are worried about being judged, some evidence suggests that they are more comfortable interacting with

an agent than a person (Pickard et al., 2016). This occurs during clinical interviews about their mental health (Slack and Van Cura, 1968; Lucas et al., 2014), but also when interviewed about their personal financial situation (Mell et al., 2017) or even during negotiations (Gratch et al., 2016). As such, the possibility exists that interactions with empathetic chatbots may be rendered safer than those with their human counterparts.

We posited that this mechanism can be adopted to help comfort participants after an experimentally induced experience of social exclusion. Many paradigms exist for inducing social exclusion, including online versions of the ball tossing game (Cyberball; Williams et al., 2000; Hartgerink et al., 2015), an online chatroom (Molden et al., 2009), and more recently a social media based social exclusion paradigm (Ostracism Online; Wolf et al., 2015). The latter simulates a social media platform such as Facebook, in which participants are excluded by receiving far fewer "likes" that other users.

Using the Ostracism Online paradigm, participants in the present work were exposed to an experience of social exclusion. Expecting to replicate prior work (e.g., Gerber and Wheeler, 2009; Lustenberger and Jagacinski, 2010; see Williams, 2007 for a review), we predicted that:

H1: Being excluded on the social media platform will negatively impact participants' mood, decreasing positive emotion and increasing negative emotion from pre- to post-exclusion.

Previous studies have established that some interventions can help excluded individuals recover from the adverse effects of social exclusion (Gross, 2009; Aydin et al., 2012). Here, we explore the possibility that an empathetic chatbot can buffer against the negative effects of social exclusion, particularly dampened mood. Therefore, our second and primary prediction is that:

H2: After being socially excluded, experiencing a conversation with an empathic chatbot will result in better mood than a comparable control experience.

To test this possibility, we created a chatbot called "Rose" to comfort participants who had just experienced social ostracism. Informed by previous research in affective computing (Picard, 2000), Rose provided empathetic responses to help them recover from the experience. To isolate the effect of such an empathetic chatbot, participants either talked about their social exclusion experiences with Rose which responded empathically (e.g., "I'm sorry that this happened to you") or took part in a control condition where their responses about the experience of social exclusion were merely acknowledged (e.g., "Thank you for your feedback"). The control condition ruled out the possibility that any differences in mood were simply due to participants disclosing personal feelings and then letting go of them. Instead, following Chaudoir and Fisher (2010), we wanted to demonstrate that the effect of talking is caused by receiving social support from the empathetic chatbot rather than through alleviating inhibition, where disclosure is beneficial merely because it allows people to express pent up emotions and thoughts (e.g., Lepore, 1997; Pennebaker, 1997).

MATERIALS AND METHODS

Participants and Design

One hundred and thirty-three participants were recruited via a department subject pool, and took part in the experiment in exchange for monetary payment. Due to technical issues, data from five participants were not usable, leaving a total of 128 (94 women; M_{age} = 24.12, SD = 5.91). A power analysis using G*Power (Faul et al., 2007) indicated that this sample size enabled approximately 80% power to detect a medium-sized effect of condition (Cohen's d = 0.50, α = 0.05, two-tailed). The study received ethical approval from the Department of Experimental Psychology, University College London.

Procedure

Informed consent was obtained prior to participation. Participants arrived individually at the laboratory, and were seated in front of a computer workstation running the Windows operating system with a 21″ screen displaying Mozilla Firefox web browser. After providing demographic information, they were told that the study was on social media profiles.

To test H1, participants were first exposed to the Ostracism Online paradigm (Wolf et al., 2015). It consists of a web-based ostracism task that has the appearance of a social media platform like Facebook and uses "likes" to either socially exclude or include participants. For the purpose of this experiment, participants were led to believe that they were going to interact with other students from the university with whom they would be connected via the internet. In reality, scripts were used to automate the experience on the platform. Minor changes were made to the original task by reducing the number of scripted "participants" from 11 to 6 to facilitate group cohesion (Bales and Borgatta, 1966; Tulin et al., 2018), and by altering some of the social media profiles to better resemble London university students.

At the beginning of the task, participants were asked to create a social media profile, i.e., choose an avatar for themselves (from among 82 options), provide a name, and write a description to introduce themselves to the other people in the group. Upon completion of the profile, they were sent to a "group page" that displayed their profile alongside those of six other students (see **Figure 1**). Participants could "like" another profile using a button that appeared under it. Whenever participants received a "like," a pop-up message notified them, indicating the name of the member who had liked their description (e.g., "Nick liked your post"). The number of "likes" was tallied at the bottom of each member's profile, which was incremented with each new "like." To induce feelings of social exclusion, the participant's profile received only one "like" while the other profiles received on average four "likes."

Experimental Conditions

To test H2, participants were then randomly assigned to one of two conditions (chatbot vs. control) in our between-subjects design. Specifically, they experienced either the empathic chatbot intervention (n = 64) or the control questionnaire (n = 64). In the chatbot intervention condition, participants interacted

FIGURE 1 | Social media exclusion induction with profile descriptions and number of "likes" for all group members. The participant's own biography always appeared on the top left corner, receiving one "like."

with a web-based embodied conversational agent named "Rose." Rose was created using the open-source platform Botkit, and was displayed as a female agent on the right side of the chat window using the open-source library GIF-Talkr (see **Figure 2**). We chose a human form based on prior work showing chatbots that have a virtual body and look human-like tend to be more effective than text alone (Koda and Maes, 1996; André et al., 1998; Khashe et al., 2017), particularly females (Kavakli et al., 2012; Khashe et al., 2017).

The user interface guided users step-by-step through the chatbot conversation (Zamora, 2017). Following the use of text- and menu-based interfaces on standard social media platforms (e.g., Facebook, 2019), participants conversed with

the chatbot mainly via multiple-choice menus (Bickmore and Schulman, 2007; Nguyen and Masthoff, 2009), which were updated in succession depending on the current conversation topic (Bickmore and Picard, 2005), and only occasionally using the text interface. Natural language interface was not used because of the risk for conversational errors due to poor speech recognition. Also, such errors have been shown to reduce trust and acceptance of agents (Blascovich and McCall, 2013; Wang et al., 2013; Lucas et al., 2018a), with chatbots attempting to use an entirely natural language interface being prone for failure (Luger and Sellen, 2016; see also Harnad, 1990). While multiple-choice menus were chosen over speech recognition, *agent* speech has been shown to increase trust and perceived
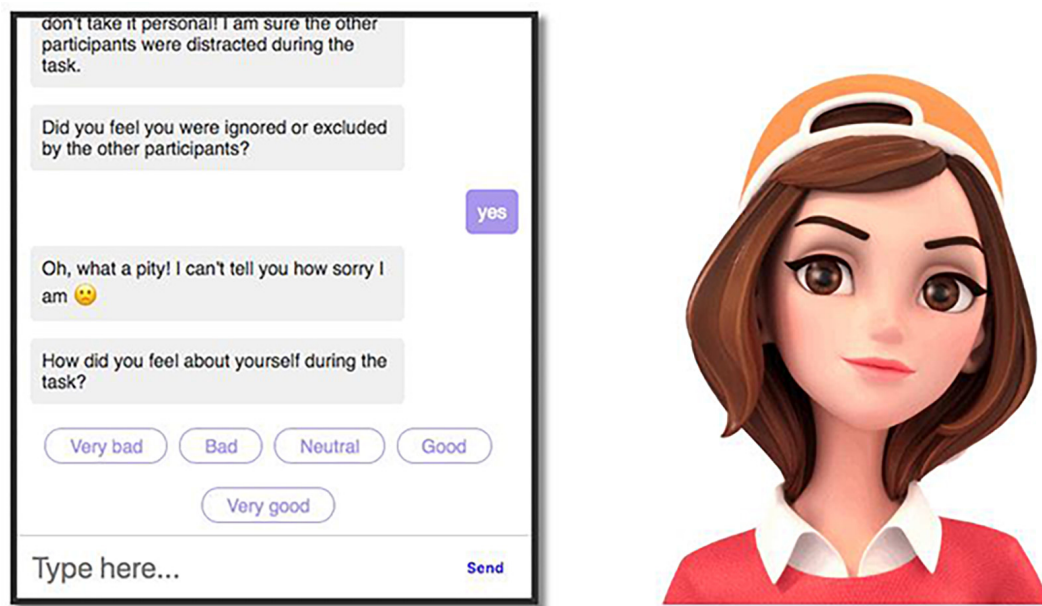
FIGURE 2 | Illustration of the chat window with the empathic chatbot "Rose."

humanness (Khashe et al., 2017; Burri, 2018; Schroeder and Schroeder, 2018). Accordingly, the chatbot used voice to interact with participants; a HTML5 Web Speech API enabled text-to-speech (TTS). Following Bhakta et al. (2014), a human-like voice (i.e., Microsoft Zira Desktop's English voices) was selected, and the agent's mouth movements were synchronized to speech produced by this TTS. To create a natural appearance the agent also blinked every 3–8 s, with the agent's responses being slightly delayed to simulate a natural flow of conversation (e.g., Zumstein and Hundertmark, 2017).

The intervention started with an informal conversation in which Rose engaged in small talk about herself to create a sense of rapport and likability (Morkes et al., 1999): "Hi! My name is Rosalind, but people typically call me Rose. I am an artificial intelligence chatbot who enjoys talking to people. I can also sing and dance, but that's probably not what you want to see right now, believe me! Anyway, I always appreciate when people try to think of me as someone with human traits, emotions, and intentions. It feels so much better to be seen as a living being." This kind of social exchange is often used as an "icebreaker" in education and organizational fields (Clear and Daniels, 2001; Risvik, 2014; Mitchell and Shepard, 2015) to build trust and rapport, which are essential for the creation of a social relationship that makes participants feel comfortable before discussing sensitive topics (Dunbar, 1998; Bickmore and Cassell, 2000). Also, this approach has been used successfully in past studies (Inaguma et al., 2016; Lucas et al., 2018a,b). The chatbot then asked a series of questions about the social media task the participant had just completed. Throughout the conversation, empathy was expressed whenever the participant reported negative feelings. For example, when someone indicated feeling excluded or ignored (Question 7), the chatbot replied: "Oh, what a pity! I can't tell you how sorry I

am." Similarly, if the participant stated that other members had not liked their social media biography (Question 6), the chatbot's answer was: "I'm sorry that this happened to you, but don't take it personal! I'm sure the other participants were distracted during the task." When participants indicated that they had not enjoyed reading the social media biographies (Question 4), the chatbot replied: "Oh, I'm sorry to hear that:(But thank you for your honesty!" Other statements were: "Thanks for telling me. BTW, I was just curious, I wasn't trying to test your memory;-)" (Question 2); "Oh, I'm sorry to hear that. I hope this conversation lifts your spirits" (Question 8); "It's fine, I was being curious, I'm sure you chose a unique avatar!" (Question 3); "Oh, I'll report it to the developers, thanks!" (Question 5). For some responses, emojis that matched the sentiment of the verbal statement were included as an additional way of providing emotional support (Smestad, 2018). In order to avoid false expectations about its capabilities (DiSalvo et al., 2002; Mori et al., 2012), the agent never explicitly used a human backstory or persona (Piccolo et al., 2018).

In contrast to the chatbot intervention condition, participants in the control condition completed an interactive questionnaire where they merely received acknowledgment that their responses were received. The number and type of questions being raised within the interface were identical to the chatbot condition. However, the conversational empathic agent was not present to provide support or comfort. For example, if a participant indicated that s/he felt ignored or excluded by the other members in the social media task (Question 7), a message like "Thank you for your feedback" was displayed to merely acknowledge receipt of their response. Similarly, if the participant stated that other members had not liked their social media biography (Question 6), the response was "Thank you for letting us know." Other statements were: "Thanks. This question was just for statistics"

(Question 2); "Thank you for your honesty" (Question 4); "We will report this to the developers, thanks" (Question 5).

Attention and Induction Check Items

Participants then completed a number of measures, attention and induction checks, followed by dependent (i.e., mood) measures. In both conditions, participants responded to eight questions (adapted from Wolf et al., 2015) about the social media task (see **Table 1**). Three questions served as an attention check to ensure that participants were appropriately engaged in the task. Another two questions assessed satisfaction with the platform. The final two questions served as an "induction check" for feelings of exclusion, as we sought to confirm that the Ostracism Online paradigm was successful in inducing feelings of exclusion.

As shown in **Table 1**, the overwhelming majority of participants passed the two most obvious attention checks by correctly answering Question 1 ("More than 4") and Question 3 ("No"). While participants were less accurate for the most challenging attention check, Question 2 (the number of participants with female names, with correct answer of "4"), given the high levels of accuracy on the other two attention checks, we take this as sufficient confirmation that participants paid attention during the experiment. On the subsequent two questions, users reported being generally satisfied with the social media platform, indicating that there were no major problems with the interface. Importantly, the Ostracism Online paradigm appeared to be effective in creating an experience of social exclusion. Most participants were aware that others did not like their profile description and the majority of participants felt at least slightly excluded. Instead of the typical skewed distribution of positive self-perception (e.g., Baumeister et al., 1989), participants showed after the ostracism task a more standard distribution for self-feelings in our study.

Measures

A different mood measure was used to assess each hypothesis. For both measures, short instruments were selected because shorter mood scales have been shown to be more sensitive to the effect of social exclusion (Gerber and Wheeler, 2009). To test the effect of social exclusion on mood (H_1), mood was measured before and after the social exclusion task using the 10-item version of the Positive and Negative Affect Scale (I-PANAS-SF; Thompson, 2007). Items assessed both positive affect (alert, inspired, determined, attentive, and active) and negative affect (upset, hostile, ashamed, nervous, and afraid), and were rated on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). Composite scores for positive and negative affect were both internally consistent (α = 0.84 and α = 0.80, respectively).

To test the moderating effect of the chatbot intervention (H_2), participants were asked to also report their mood after the chatbot (or control) experience. In order to prevent participants from noticing that the scale was the same, and thus allow demand characteristics to influence responses, final mood was assessed using the question "How do you feel at this particular moment?" (1 = *extremely negative*, 7 = *extremely positive*). At the end of the study, participants were debriefed, thanked and compensated for their time. Prior to revealing the purpose of the experiment

during the debriefing, no participants indicated suspicions about any aspects of the study.

RESULTS

To test the effect of social exclusion on mood (H_1), we compared participants' affect ratings before and after the Ostracism Online experience. Composite scores of the I-PANAS-SF were submitted to two paired-samples t tests with positive or negative mood as the dependent measure. After being socially excluded, positive affect was reduced (M_{pre} = 16.80, SD = 3.45 vs. M_{post} = 13.97, SD = 4.46), $t(127)$ = 9.08, p < 0.001, Cohen's d = 0.80, 95% CI [0.46,0.96], whereas negative affect increased (M_{pre} = 6.89, SD = 2.63 vs. M_{post} = 8.33, SD = 3.42), $t(127)$ = -4.70, p < 0.001, Cohen's d = 0.42, 95% CI [0.22,0.72]. Interpreting the confidence intervals around the effect sizes, with 95% confidence, the effect is small to large for the increase in negative affect and medium to large for the decrease in positive affect. The same results were obtained when submitting the data to a 2 (time: pre vs. post) x 2 (valence: positive vs. negative) repeated-measures ANOVA, showing significant main effects of time, $F(1,127)$ = 11.64, p = 0.001, η_p^2 = 0.08, and valence, $F(1,127)$ = 355.37, p < 0.001, η_p^2 = 0.74, as well as their interaction, $F(1,127)$ = 85.13, p < 0.001, η_p^2 = 0.40. None of the above effects was moderated by condition (chatbot vs. control), Fs < 1.03, ps > 0.31, indicating that mood changed as a function of the Ostracism Online experience similarly across both conditions.

To test the effect of the chatbot intervention (H_2), an independent samples t-test was conducted on mood after the intervention (or control) experience. In line with predictions, a main effect of condition occurred, $t(126)$ = 2.08, p = 0.040, Cohen's d = 0.37, 95% CI [0.02,0.72]; interpreting the confidence interval around the effect size, with 95% confidence, the effect of condition is non-negative to large. After being socially excluded, those who interacted with the chatbot had significantly more positive mood (M = 3.77, SD = 0.89) compared to participants in the control condition (M = 3.45, SD = 0.82). When dropping data from participants who responded feeling "not at all" socially excluded (Question 7), the results remained the same, $M_{chatbot}$ = 3.62, SD = 0.79, $M_{control}$ = 3.30, SD = 0.72, $t(92)$ = 2.04, p = 0.044, Cohen's d = 0.42, 95% CI [0.13,0.71]; interpreting the confidence interval around the effect size, with 95% confidence, the effect is small to large. When considering feelings of social exclusion (Question 7) as a covariate in an ANCOVA, the main effect of condition remained significant, $M_{chatbot}$ = 3.77, SD = 0.89, $M_{control}$ = 3.45, SD = 0.82, $F(1,125)$ = 4.72, p = 0.032, η_p^2 = 0.04. We are therefore confident that the results are stable, even in the context of those participants who failed the induction check for feelings of exclusion.

DISCUSSION

In this work, we provide initial evidence that a fully automated embodied empathetic agent has the potential to improve users' mood after experiences of social exclusion. First, we expected

TABLE 1 | Response frequencies for the eight questions about the social media task.

1. How many participants were connected to the social media platform?	2	3	4	More than 4	
	0.78% (N = 1)	0.78% (N = 1)	0.78% (N = 1)	97.66% (N = 125)	
2. Do you remember how many participants had female names?	Unsure	2	3	4	More than 4
	10.16% (N = 13)	5.47% (N = 7)	32.03% (N = 41)	32.81% (N = 42)	19.53% (N = 25)
3. Did any of the participants choose the same avatar as you did?	Yes	No	Unsure		
	1.56% (N = 2)	93.75% (N = 120)	4.69% (N = 6)		
4. Have you enjoyed reading the social media biographies?	Yes	No			
	82.81% (N = 106)	17.19% (N = 22)			
5. Were you able to "like" other people's posts and see their and your "likes"?	Yes	No			
	96.09% (N = 123)	3.91% (N = 5)			
6. Do you think the other participants liked your description?	Yes	No			
	30.47% (N = 39)	69.53% (N = 89)			
7. Did you feel you were ignored or excluded by the other participants?	Not at all	Slightly Excluded	Excluded	Extremely Excluded	
	26.56% (N = 34)	56.25% (N = 72)	15.63% (N = 20)	1.56% (N = 2)	
8. How did you feel about yourself during the task?	Very Bad	Bad	Neutral	Good	Very Good
	0.78% (N = 1)	12.5% (N = 16)	57.81% (N = 74)	19.53% (N = 25)	9.38% (N = 12)

to replicate previous findings that being excluded would have a negative impact on mood. Indeed, a comparison of the I-PANAS-SF scores before and after the social exclusion task revealed a significant increase in negative mood and a significant decline in positive mood. Although the effect of social exclusion on subsequent behavior may last longer than its effect on mood (Williams, 2007; Gerber and Wheeler, 2009), research has shown substantial immediate effects on mood, with exclusion lowering positive affect and increasing negative affect (see Williams, 2007 for a review).

Importantly, in our second hypothesis, we predicted that an emotional support chatbot that displays empathy would mitigate the negative effect on mood. As expected, the chatbot intervention helped participants to have a more positive mood (compared to the control condition) after being socially excluded. This result is in line with those of previous studies with emotional support chatbots designed for other purposes (e.g. Bickmore and Picard, 2005; Nguyen and Masthoff, 2009), and further

supports the idea that chatbots that display empathy may have the potential to help humans recover more quickly after experiencing social ostracism.

A possible explanation for this finding could be offered by Media Equation Theory (Reeves and Nass, 1996), which states that humans instinctively perceive and react to computers (and other media) in much the same manner as they do with people. Despite knowing that computers are inanimate, there is evidence that they unconsciously attribute human characteristics to computers and treat them as social actors (Nass and Moon, 2000). Moreover, people often rely on heuristics or cognitive shortcuts (Tversky and Kahneman, 1974) and mindlessly apply social scripts from human-human interaction when interacting with computers (Sundar and Nass, 2000). Nass and Moon (2000) argue that we tend not to differentiate mediated experiences from non-mediated experiences and focus on the social cues provided by machines, effectively "suspending disbelief" in their humanness. Due to our social nature, we may fail to distinguish

chatting with a bot from interacting with a fellow human. As such, there is reason to believe that people have a strong tendency to respond to the social and emotional cues expressed by the chatbot in a way as if they had originated from another person. For example, Liu and Sundar (2018) found evidence supporting the Media Equation Theory in the context of chatbots expressing sympathy, cognitive empathy and affective empathy. In line with this notion, sympathy or empathy coming from a chatbot could then have similar effects on the individual as in human-human interaction.

The present research tried to rule out the possibility that the observed differences in mood between the chatbot intervention and control conditions were due to participants disclosing about, and thus letting go of, the social exclusion. Given that participants in both conditions had the opportunity to get the exclusion experience "off their chest," the control condition was equated to the intervention condition in terms of alleviating inhibition (Chaudoir and Fisher, 2010), allowing people to express pent up emotions and thoughts (e.g., Lepore, 1997; Pennebaker, 1997). We are therefore relatively confident that mood was restored through the provision of social support by the empathic chatbot rather than just letting users express themselves.

The present research makes important contributions. First and foremost, the work appears to be the first to evaluate the usefulness of chatbots in helping individuals deal with the negative effects of social exclusion. As such, it demonstrates the possibility of empathic chatbots as a supportive technology in the face of social exclusion. Additionally, by showing that empathetic chatbots have the potential to recover mood after exclusion on social media, the work contributes to both the social exclusion literature and the field of human-computer interaction. By adapting the Ostracism Online task (Wolf et al., 2015) for the purposes of the present research, we validated the paradigm in a different setting (i.e., laboratory) with university students rather than online via Mechanical Turk workers (see support for H_1). Furthermore, it extends most past studies in human-computer interaction which used the Wizard of Oz methodology (see Dahlbäck et al., 1993) in which participants are led to believe that they are interacting with a chatbot when in fact the chatbot is being remotely controlled by a human confederate. The present study employed a fully automated empathic chatbot. Since this chatbot was created using free open source tools, it can be easily customized for future research or even be of applied use to health professionals. This makes a final contribution by affording opportunities for future research and applications.

Limitations and Further Research

Before drawing conclusions regarding the effectiveness of empathic chatbots in assisting socially excluded individuals, it is essential to examine whether novelty effects contributed to the results. According to the innovation hypothesis, any social reaction toward chatbots is simply due to novelty which eventually disappears once the novelty wears off (Chen et al., 2016; Fryer et al., 2017). In future research, longitudinal studies could be conducted in order to rule out this possibility. Moreover, similar but sufficiently distinct mood scales could be used over the course of the experiment to allow for direct comparability in

mood between the different time points. In the present research, a single item affect measure was employed at time 3 to prevent participants from indicating the same response several times. While such approach avoids potential demand effects, it did not allow us to measure direct change in affect from time 2 to 3. Alternatively, rather than relying on self-report scales, future studies might consider implicit measures of mood.

Another limitation was the limited number of participants that were recruited, and the relatively high p-value for the main effect of condition (H2). The small number of participants meant that only two conditions could be studied (chatbot vs. control), as a 2 (exclusion vs. inclusion) x 2 (chatbot vs. control) design would have resulted in insufficient statistical power. To address this shortcoming, future research should replicate this study with more participants, adding a social inclusion condition to the design. While the control condition was otherwise perfectly comparable to the chatbot intervention condition, it is important to note that the experimental design does manipulate both empathy and the presence of a chatbot. Thus, in this initial test we considered the "empathetic chatbot" as the intervention under investigation. Furthermore, the current study's control condition differed from the intervention condition in terms of the presence of empathy and the chatbot itself. Accordingly, future research could further isolate "empathy" as the driving factor (controlling the mere presence of the chatbot) by employing a control condition with a chatbot that does not attempt to make participants feel better. While the resulting effect would likely be smaller (and thus require a larger sample size to achieve comparable statistical power), this design would increase internal validity.

Likewise, the observed effect may have been bolstered by the presence of a human-like face (compared to no face). For example, there is evidence that people perceive embodied chatbots that look like humans as more empathic and supportive than otherwise equivalent chatbots that are not embodied (i.e., text-only; Nguyen and Masthoff, 2009; Khashe et al., 2017). Future research could consider the role of embodiment by comparing the effectiveness of *embodied* empathetic chatbots for ameliorating negative effects of social exclusion to the effectiveness of equivalent chatbots that are not embodied.

While this research suggests that chatbots can help humans recover their mood more quickly after social exclusion, empathetic agents may reduce the willingness to seek social connection, especially for lonely individuals given that they fear social rejection (Lucas, 2010; Lucas et al., 2010). For example, work on "social snacking" demonstrates that social cues of acceptance (such as reading a message from a loved one) can temporarily satiate social needs and in turn reduce attempts to connect (Gardner et al., 2005). Accordingly, it is possible that agents that build connection using empathy and other rapport-building techniques could cue social acceptance, thereby lowering users' willingness to reach out to others. Krämer et al. (2018) provided initial evidence for this possibility by demonstrating that, among those with activated needs to belong (i.e., lonely or socially isolated individuals), users were less willing to try to connect with other humans after interacting with a virtual agent.

This occurred only if the agent displayed empathetic, rapport-building behavior. By meeting any outstanding immediate social needs, empathetic chatbots could therefore make users more socially apathetic. Over the long term, this might hamper people from fully meeting their need to belong. If empathetic chatbots draw us away from real social connection with other humans through a fleeting sense of satisfaction, there is an especially concerning risk for those who suffer from chronic loneliness, given they are already hesitant to reach out to others so as not to risk being rejected. As such, supportive social agents, which are perceived as safe because they will not negatively evaluate or reject them (Lucas et al., 2014), could be very alluring to people with chronic loneliness, social anxiety, or otherwise heightened fears of social exclusion. But those individuals, who already feel disconnected, are likely to not find their need to belong truly fulfilled by these "parasocial" interactions. Future research should thus consider these possibilities and seek to determine under what conditions -and for whom- empathetic chatbots are able to encourage attempts at social connection.

Finally, while this research suggests that chatbots can help humans recover their mood more quickly after social exclusion, this intervention would not serve as the sole remedy for the effect of social exclusion on mood and mental health. While intense interventions such as cognitive-behavioral therapy, acceptance and commitment therapy, and dialectical behavioral therapy can help people to learn how to cope with the negative feelings and reframe the rejection such that it does not produce such negative affect and adverse effects on mental health, there are also other simpler more subtle interventions that could be used -like empathic chatbots- to reduce the sting of rejection and its impact on mood. For example, results from Lucas et al. (2010) suggest that subtly priming social acceptance may be able to trigger "upward spiral" of positive reaction and mood among those faced with perceived rejection; this suggests that "even the smallest promise of social riches" can begin to ameliorate the negative impact of rejection.

CONCLUSION

Adding to the literature on how to achieve social impact with chatbots, this study yields promising evidence that ECAs have the potential to provide emotional support to victims of social exclusion. Fully automated empathic chatbots that can comfort individuals have important applications in healthcare.

In particular, they offer unique benefits such as the ability to instantly reach large amounts of people, being continuously available, and overcoming geographical barriers to care. Even if chatbots do not infiltrate healthcare, they may be effective at mitigating negative emotional effects such as those created by cyberbullying. In this and similar use cases and applications, chatbots can be deployed to support mood when users embark in the murky waters of the internet with its potential risks of negativity and hurt feelings. In such cases, empathetic chatbots should be used alongside other approaches to improve the mental health of individuals who are victims of cyberbullying. Finally, while the present results are preliminary and need to be viewed with caution, our study demonstrates the potential of chatbots as a supportive technology and sets a clear roadmap for future research.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Department of Experimental Psychology, University College London. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MG and EK conceived and designed the experiments and analyzed the data. MG performed the experiments. GL, EK, and MG contributed to writing and revising the manuscript and read and approved the submitted version.

FUNDING

REFERENCES

Ali, S. M., Merlo, J., Rosvall, M., Lithman, T., and Lindström, M. (2006). Social capital, the miniaturisation of community, traditionalism and first time acute myocardial infarction: a prospective cohort study in southern Sweden. *Soc. Sci. Med.* 63, 2204–2217. doi: 10.1016/J.SOCSCIMED.2006.04.007

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., et al. (2013). Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. *J. Consult. Clin. Psychol.* 81, 751–760. doi: 10.1037/a0033559

André, E., Rist, T., and Müller, J. (1998). "Integrating reactive and scripted behaviors in a life-like presentation agent," in *Proceedings of the Second International Conference on Autonomous Agents - AGENTS '98*, New York, NY.

Aydin, N., Krueger, J. I., Fischer, J., Hahn, D., Kastenmüller, A., Frey, D., et al. (2012). Man's best friend: how the presence of a dog reduces mental distress after social exclusion. *J. Exp. Soc. Psychol.* 48, 446–449. doi: 10.1016/J.JESP.2011.09.011

Bales, R. F., and Borgatta, E. (1966). "Size of group as a factor in the interaction profile," in *Small Groups: Studies in Social Interaction*, eds A. P. Hare, E. Borgatta, and R. Bales, (New York, NY: Knopf), 495–512.

Batrinca, L., Stratou, G., Shapiro, A., Morency, L. P., and Scherer, S. (2013). "Cicero - Towards a multimodal virtual audience platform for public speaking training," in *Intelligent Virtual Agents. IVA 2013. Lecture Notes in Computer Science*, eds R. Aylett, B. Krenn, C. Pelachaud, and H. Shimodaira, (Berlin: Springer).

Baumeister, R. F., and Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497–529. doi: 10.1037/0033-2909.117.3.497

Baumeister, R. F., and Tice, D. M. (1990). Point-counterpoints: anxiety and social exclusion. *J. Soc. Clin. Psychol.* 9, 165–195. doi: 10.1521/jscp.1990.9.2.165

Baumeister, R. F., Tice, D. M., and Hutton, D. G. (1989). Self-presentational motivations and personality differences in self-esteem. *J. Pers.* 57, 547–579. doi: 10.1111/j.1467-6494.1989.tb02384.x

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educ. Technol. Res. Dev.* 59, 291–300. doi: 10.1007/s11423-011-9196-3

Bhakta, R., Savin-Baden, M., and Tombs, G. (2014). "Sharing secrets with robots?," in *Proceedings of EdMedia + Innovate Learning 2014*, Anchorage, AK.

Bickmore, T., Asadi, R., Ehyaei, A., Fell, H., Henault, L., Intille, S., et al. (2015). "Context-awareness in a persistent hospital companion agent," in *Proceedings of the International Conference on Intelligent Virtual Agents*, eds W. P. Brinkman, J. Broekens, and D. Heylen, (Cham: Springer), 332–342. doi: 10.1007/978-3-319-21996-7_35

Bickmore, T., and Cassell, J. (2000). "How about this weather? Social dialogue with embodied conversational agents," in *Proceedings of the Socially Intelligent Agents: The Human in the Loop, Papers from the 2000 AAAI Fall Symposium*, ed. K. Dautenhahn, (Cambridge MA: Cambridge University).

Bickmore, T., and Fernando, R. (2009). "Towards empathic touch by relational agents," in *Proceedings of the Autonomous Agents and Multiagent Systems (AAMAS) Workshop on Empathic Agents*, Cambridge.

Bickmore, T., Gruber, A., and Picard, R. (2005). Establishing the computer–patient working alliance in automated health behavior change interventions. *Patient Educ. Couns.* 59, 21–30. doi: 10.1016/J.PEC.2004.09.008

Bickmore, T., and Schulman, D. (2007). "Practical approaches to comforting users with relational agents," in *Proceedings of the CHI '07 Extended Abstracts on Human Factors in Computing Systems*, ed. M. B. Rosson, (New York, NY: ACM Press), 2291.

Bickmore, T. W., and Picard, R. W. (2004). "Towards caring machines," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI' 04)*, Vienna.

Bickmore, T. W. (2003). *Relational agents: Effecting Change through Human-Computer Relationships*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Bickmore, T. W., and Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Trans. Comput. Hum. Int.* 12, 293–327. doi: 10.1145/1067860.1067867

Blascovich, J., and McCall, C. (2013). "Social influence in virtual environments," in *The Oxford Handbook of Media Psychology*, ed. K. E. Dill, (Oxford: Oxford University Press), 305–315.

Brandtzaeg, P. B., and Følstad, A. (2018). Chatbots: changing user needs and motivations. *Interactions* 25, 38–43. doi: 10.1145/3236669

Brave, S., and Nass, C. (2007). "Emotion in human-computer interaction," in *The Human-Computer Interaction Handbook*, eds A. Sears, and J. A. Jacko, (Boca Raton, FL: CRC Press), 103–118.

Brave, S., Nass, C., and Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *Int. J. Hum. Comput. Stud.* 62, 161–178. doi: 10.1016/J.IJHCS.2004.11.002

Burleson, B. R., and Goldsmith, D. J. (1996). How the comforting process works: Alleviating emotional distress through conversationally induced reappraisals. *Handb. Commun. Emot.* 1996, 245–280. doi: 10.1016/B978-012057770-5/50011-4

Burri, R. (2018). *Improving User Trust Towards Conversational Chatbot Interfaces With Voice Output (Dissertation)*. Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-240585 (accessed February 13, 2019).

Chaudoir, S. R., and Fisher, J. D. (2010). The disclosure processes model: understanding disclosure decision making and postdisclosure outcomes among people living with a concealable stigmatized identity. *Psychol. Bull.* 136:236. doi: 10.1037/a0018193

Chen, J. A., Tutwiler, M. S., Metcalf, S. J., Kamarainen, A., Grotzer, T., and Dede, C. (2016). A multi-user virtual environment to support students' self-efficacy and interest in science: a latent growth model analysis. *Learn. Instruct.* 41, 11–22. doi: 10.1016/J.LEARNINSTRUC.2015.09.007

Clear, T., and Daniels, M. (2001). "A cyber-icebreaker for an effective virtual group?," in *Proceedings of the 6th Annual Conference on Innovation and Technology in Computer Science Education*, Canterbury.

Cuijpers, P., Marks, I. M., van Straten, A., Cavanagh, K., Gega, L., and Andersson, G. (2009). Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. *Cogn. Behav. Ther.* 38, 66–82. doi: 10.1080/16506070802694776

Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993). Wizard of oz studies – why and how. *Knowl. Based Syst.* 6, 258–266. doi: 10.1016/0950-7051(93)90017-N

Davis, J. A., and Smith, T. W. (1998). *General Social Surveys, 1972-1998: Cumulative Codebook*. Chicago: NORC.

de Melo, C. M., Gratch, J., and Carnevale, P. J. (2014). Humans versus computers: impact of emotion expressions on people's decision making. *IEEE Trans. Affect. Comput.* 6, 127–136. doi: 10.1109/taffc.2014.2332471

DeVault, D., Artstein, R., Benn, G., Dey, T., Egan, A., Fast, E., et al. (2014). "SimInterviewer: a virtual human interviewer for healthcare decision support," in *Proceedings of the 13th Conference on Autonomous Agents and Multi-Agent Systems*, New York, NY.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., and Kiesler, S. (2002). "All robots are not created equal," in *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, New York, NY.

Dunbar, R. (1998). *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Eisenberger, N. I., and Lieberman, M. D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends Cogn. Sci.* 8, 294–300. doi: 10.1016/J.TICS.2004.05.010

Eisenberger, N. I., Lieberman, M. D., and Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292. doi: 10.1126/science.1089134

Epley, N., Waytz, A., and Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychol. Rev.* 114:864. doi: 10.1037/0033-295X.114.4.864

Facebook (2019). *General Best Practices - Messenger Platform*. Available at: https://developers.facebook.com/docs/messenger-platform/introduction/general-best-practices (accessed February 13, 2019)

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

Faulkner, S., Williams, K., Sherman, B., and Williams, E. (1997). "The "silent treatment": Its incidence and impact," in *Proceedings of the 69th Annual Midwestern Psychological Association*, Chicago.

Fitzpatrick, K. K., Darcy, A., and Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR Ment. Health* 4:e19. doi: 10.2196/mental.7785

Følstad, A., and Brandtzaeg, P. B. (2017). Chatbots and the new world of HCI. *Interactions* 24, 38–42. doi: 10.1145/3085558

Friedman, B. (1997). *Human Values and the Design of Computer Technology*. Cambridge: Cambridge University Press.

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., and Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: an experimental comparison of Chatbot and Human task partners. *Comput. Hum. Behav.* 75, 461–468. doi: 10.1016/J.CHB.2017.05.045

Gardner, W. L., Pickett, C. L., and Knowles, M. (2005). "Social Snacking and Shielding: Using social symbols, selves, and surrogates in the service of belonging needs," in *The Social Outcast: Ostracism, Social Exclusion, Rejection, and Bullying*, eds K. D. Williams, J. P. Forgas, and W. von Hippel, (New York, NY: Psychology Press).

Gerber, J., and Wheeler, L. (2009). On being rejected: a meta-analysis of experimental research on rejection. *Perspect. Psychol. Sci.* 4, 468–488. doi: 10.1111/j.1745-6924.2009.01158.x

Gerin, W., Pieper, C., Levy, R., and Pickering, T. G. (1992). Social support in social interaction: a moderator of cardiovascular reactivity. *Psychosom. Med.* 54, 324–336. doi: 10.1097/00006842-199205000-00008

Gonsalkorale, K., and Williams, K. D. (2007). The KKK won't let me play: ostracism even by a despised outgroup hurts. *Eur. J. Soc. Psychol.* 37, 1176–1186. doi: 10.1002/ejsp.392

Gratch, J., DeVault, D., and Lucas, G. M. (2016). "The benefits of virtual humans for teaching negotiation," in *Proceedings of the 16th International Conference on Intelligent Virtual Agent*, Berlin.

Gross, E. F. (2009). Logging on, bouncing back: an experimental investigation of online communication following social exclusion. *Dev. Psychol.* 45, 1787–1793. doi: 10.1037/a0016541

Harnad, S. (1990). The symbol grounding problem. *Physica D* 42, 335–346. doi: 10.1016/0167-2789(90)90087-6

Hartgerink, C. H. J., van Beest, I., Wicherts, J. M., and Williams, K. D. (2015). The ordinal effects of ostracism: a meta-analysis of 120 cyberball studies. *PLoS One* 10:e0127002. doi: 10.1371/journal.pone.0127002

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

Hone, K. (2006). Empathic agents to reduce user frustration: the effects of varying agent characteristics. *Interact. Comput.* 18, 227–245. doi: 10.1016/j.intcom.2005.05.003

Huang, J., Li, Q., Xue, Y., Cheng, T., Xu, S., Jia, J., et al. (2015). "TeenChat: a chatterbot system for sensing and releasing adolescents' stress," in *International Conference on Health Information Science*, eds X. Yin, K. Ho, D. Zeng, U. Aickelin, R. Zhou, and H. Wang, (Cham: Springer), 133–145. doi: 10.1007/978-3-319-19156-0_14

IJzerman, H., Gallucci, M., Pouw, W. T., Weißgerber, S. C., Van Doesum, N., and Williams, K. D. (2012). Cold-blooded loneliness: social exclusion leads to lower skin temperatures. *Acta Psychol.* 140, 283–288. doi: 10.1016/J.ACTPSY.2012.05.002

Inaguma, H., Inoue, K., Nakamura, S., Takanashi, K., and Kawahara, T. (2016). "Prediction of ice-breaking between participants using prosodic features in the first meeting dialogue," in *Proceedings of the 2nd Workshop on Advancements in Social Signal Processing for Multimodal Interaction*, New York, NY.

Jarrell, P., Brian, A., Matthieu, D., Anton, T., Matt, L., Ken, G., et al. (2006). "A virtual reality exposure therapy application for Iraq war post-traumatic stress disorder," in *Proceedings of the IEEE Virtual Reality Conference (VR 2006)*, Piscataway, NJ.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Acad. Manag. J.* 33, 692–724. doi: 10.5465/256287

Kavakli, M., Li, M., and Rudra, T. (2012). Towards the development of a virtual counselor to tackle students' exam stress. *J. Integr. Des. Process Sci.* 16, 5–26. doi: 10.3233/JID-2012-0004

Kawamoto, T., Nittono, H., and Ura, M. (2013). Cognitive, affective, and motivational changes during ostracism: an ERP, EMG, and EEG study using a computerized cyberball task. *Neurosci. J.* 2013:304674. doi: 10.1155/2013/304674

Khashe, S., Lucas, G. M., Becerik-Gerber, B., and Gratch, J. (2017). Buildings with persona: towards effective building-occupant communication. *Comput. Hum. Behav.* 75, 607–618. doi: 10.1016/j.chb.2017.05.040

Kim, J., Kim, Y., Kim, B., Yun, S., Kim, M., and Lee, J. (2018). "Can a machine tend to teenagers' emotional needs?," in *Proceedings of the Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, New York, NY.

Klein, J., Moon, Y., and Picard, R. W. (2002). This computer responds to user frustration. *Interact. Comput.* 14, 119–140. doi: 10.1016/S0953-5438(01)00053-4

Koda, T., and Maes, P. (1996). "Agents with faces: the effect of personification," in *Proceedings 5th IEEE International Workshop on Robot and Human Communication. RO-MAN'96 TSUKUBA*, Piscataway, NJ.

Krämer, N., Lucas, G. M., Schmitt, L., and Gratch, J. (2018). Social snacking with a virtual agent – On the interrelation of need to belong and social effects of rapport when interacting with artificial entities. *Int. J. Hum. Comput. Stud.* 109, 112–121. doi: 10.1016/j.ijhcs.2017.09.001

Krämer, N. C., von, der Pütten, A., and Eimler, S. (2012). "Human-agent and human-robot interaction theory: similarities to and differences from human-human interaction," in *Human-Computer Interaction: The Agency Perspective*, eds M. Zacarias, and J. V. Oliveira, (Berlin: Springer), 215–240. doi: 10.1007/978-3-642-25691-2-9

Lancaster, J. B. (1986). Primate social behavior and ostracism. *Ethol. Sociobiol.* 7, 215–225. doi: 10.1016/0162-3095(86)90049-X

Leary, M. R. (1990). Responses to social exclusion: social anxiety, jealousy, loneliness, depression, and low self-esteem. *J. Soc. Clin. Psychol.* 9, 221–229. doi: 10.1521/jscp.1990.9.2.221

Leary, M. R., Kowalski, R. M., Smith, L., and Phillips, S. (2003). Teasing, rejection, and violence: case studies of the school shootings. *Aggres. Behav.* 29, 202–214. doi: 10.1002/ab.10061

Leary, M. R., Tambor, E. S., Terdal, S. K., and Downs, D. L. (1995). Self-esteem as an interpersonal monitor: the sociometer hypothesis. *J. Pers. Soc. Psychol.* 68:518. doi: 10.1037//0022-3514.68.3.518

Lepore, S. J. (1997). Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. *J. Pers. Soc. Psychol.* 73, 1030–1037. doi: 10.1037/0022-3514.73.5.1030

Lim, D. S. (2017). *Predicting Outcomes in Online Chatbot-Mediated Therapy*. Available at: http://cs229.stanford.edu/proj2017/final-reports/5231013.pdf (accessed February 13, 2019).

Liu, B., and Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychol. Behav. Soc. Netw.* 21, 625–636. doi: 10.1089/cyber.2018.0110

Lucas, G. M. (2010). Wanting connection, fearing rejection: the paradox of loneliness. *Human* 17, 37–41.

Lucas, G. M., Boberg, J., Artstein, R., Traum, D., Gratch, J., Gainer, A., et al. (2018a). "Culture, errors, and rapport-building dialogue in social agents," in *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, Los Angeles, CA.

Lucas, G. M., Boberg, J., Artstein, R., Traum, D., Gratch, J., Gainer, A., et al. (2018b). "Getting to know each other: the role of social dialogue in recovery from errors in social robots," in *Proceedings of the Conference on Human-Robot Interaction*, Chicago.

Lucas, G. M., Gratch, J., King, A., and Morency, L.-P. (2014). It's only a computer: virtual humans increase willingness to disclose. *Comput. Hum. Behav.* 37, 94–100. doi: 10.1016/J.CHB.2014.04.043

Lucas, G. M., Knowles, M. L., Gardner, W. L., Molden, D. C., and Jefferis, V. E. (2010). Increasing social engagement among lonely individuals: the role of acceptance cues and promotion motivations. *Pers. Soc. Psychol. Bull.* 36, 1346–1359. doi: 10.1177/0146167210382662

Lucas, G. M., Rizzo, A. S., Gratch, J., Scherer, S., Stratou, G., Boberg, J., et al. (2017). Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. *Front. Robot.* 2017:4. doi: 10.3389/frobt.2017.00051

Luger, E., and Sellen, A. (2016). "Like having a really bad PA": the gulf between user expectation and experience of conversational agents," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, New York, NY.

Lustenberger, D. E., and Jagacinski, C. M. (2010). Exploring the effects of ostracism on performance and intrinsic motivation. *Hum. Perform.* 23, 283–304. doi: 10.1080/08959285.2010.501046

Mell, J., Lucas, G. M., and Gratch, J. (2017). "Prestige questions, online agents, and gender-driven differences in disclosure," in *Intelligent Virtual Agents. IVA 2017. Lecture Notes in Computer Science*, eds J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, B. Leite, and S. Kopp, (Cham: Springer).

Meyer, J., Miller, C., Hancock, P., de Visser, E. J., and Dorneich, M. (2016). Politeness in machine-human and human-human interaction. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* 60, 279–283. doi: 10.1177/1541931213601064

Miner, A., Chow, A., Adler, S., Zaitsev, I., Tero, P., Darcy, A., et al. (2016). "Conversational agents and mental health," in *Proceedings of the Fourth International Conference on Human Agent Interaction - HAI '16*, New York, NY.

Mitchell, S. K., and Shepard, M. (2015). "Building social presence through engaging online instructional strategies," in *Student-Teacher Interaction in Online Learning Environments*, ed. R. D. Wright, (Hershey, PA: IGI Global), 133–156. doi: 10.4018/978-1-4666-6461-6.ch007

Molden, D. C., Lucas, G. M., Gardner, W. L., Dean, K., and Knowles, M. L. (2009). Motivations for prevention or promotion following social exclusion: being rejected versus being ignored. *J. Pers. Soc. Psychol.* 96, 415–431. doi: 10.1037/a0012958

Mori, M., MacDorman, K., and Kageki, N. (2012). The uncanny valley. *IEEE Robot. Autom. Magaz.* 19, 98–100. doi: 10.1109/MRA.2012.2192811

Morkes, J., Kernal, H. K., and Nass, C. (1999). Effects of humor in task-oriented human-computer interaction and computer-mediated communication: a direct test of SRCT theory. *Hum. Comput. Interact.* 14, 395–435. doi: 10.1207/S15327051HCI1404_2

Nass, C., and Moon, Y. (2000). Machines and mindlessness: social responses to computers. *J. Soc. Issues* 56, 81–103. doi: 10.1111/0022-4537.00153

Nass, C., Moon, Y., and Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *J. Appl. Soc. Psychol.* 27, 864–876. doi: 10.1111/j.1559-1816.1997.tb00275.x

Nguyen, H., and Masthoff, J. (2009). "Designing empathic computers," in *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*, New York, NY.

Nolan, S. A., Flynn, C., and Garber, J. (2003). Prospective relations between rejection and depression in young adolescents. *J. Pers. Soc. Psychol.* 85, 745–755. doi: 10.1037/0022-3514.85.4.745

Parise, S., Kiesler, S., Sproull, L., and Waters, K. (1999). Cooperating with life-like interface agents. *Comput. Hum. Behav.* 15, 123–142. doi: 10.1016/S0747-5632(98)00035-1

Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychol. Sci.* 8, 162–166. doi: 10.1111/j.1467-9280.1997.tb00403.x

Peplau, L. A., and Perlman, D. (1982). *Loneliness: A Sourcebook of Current Theory, Research, and Therapy*. New York, NY: Wiley.

Petersen, S., Houston, S., Qin, H., Tague, C., and Studley, J. (2017). The utilization of robotic pets in dementia care. *J. Alzheimer Dis.* 55, 569–574. doi: 10.3233/jad-160703

Picard, R. W. (2000). *Affective Computing*. Cambridge, MA: MIT Press.

Piccolo, L. S. G., Roberts, S., Iosif, A., and Alani, H. (2018). "Designing chatbots for crises: a case study contrasting potential and reality," in *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI)* (New York, NY: ACM Press), 1–10.

Pickard, M. D., Roster, C. A., and Chen, Y. (2016). Revealing sensitive information in personal interviews: is self-disclosure easier with humans or avatars and under what conditions? *Comput. Hum. Behav.* 65, 23–30. doi: 10.1016/j.chb.2016.08.004

Pickett, C. L., and Gardner, W. L. (2005). "The social monitoring system: enhanced sensitivity to social cues as an adaptive response to social exclusion," in *The social Outcast: Ostracism, Social Exclusion, Rejection, and Bullying*, eds K. D. Williams, J. P. Forgas, and W. von Hippel, (New York, NY: Psychology Press), 213–226.

Prendinger, H., Mori, J., and Ishizuka, M. (2005). Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *Int. J. Hum. Comput. Stud.* 62, 231–245. doi: 10.1016/J.IJHCS.2004.11.009

Provoost, S., Lau, H. M., Ruwaard, J., and Riper, H. (2017). Embodied conversational agents in clinical psychology: a scoping review. *J. Med. Internet Res.* 19:e151. doi: 10.2196/jmir.6553

Reblin, M., and Uchino, B. N. (2008). Social and emotional support and its implication for health. *Curr. Opin. Psychiatr.* 21, 201–205. doi: 10.1097/YCO.0b013e3282f3ad89

Reeves, B., and Nass, C. I. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People And Places*. Cambridge MA: Cambridge University Press.

Reynolds, W. J. (2017). *The Measurement and Development of Empathy in Nursing*. Abingdon: Routledge.

Risvik, L. L. (2014). *Framework for the Design and Evaluation of Icebreakers in Group Work*. Master's Thesis, University of Oslo, Oslo.

Rizzo, A., Buckwalter, J. G., John, B., Newman, B., Parsons, T., Kenny, P., et al. (2012). STRIVE: stress resilience in virtual environments: a pre-deployment VR system for training emotional coping skills and assessing chronic and acute stress responses. *Stud. Health Technol. Inform.* 173, 379–385.

Schroeder, J., and Schroeder, M. (2018). "Trusting in machines: how mode of interaction affects willingness to share personal information with machines," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, Honolulu, HI.

Silk, J. B., Alberts, S. C., and Altmann, J. (2003). Social bonds of female baboons enhance infant survival. *Science* 302, 1231–1234. doi: 10.1126/science.1088580

Slack, W. V., and Van Cura, L. J. (1968). Patient reaction to computer-based medical interviewing. *Comput. Biomed. Res.* 1, 527–531. doi: 10.1016/0010-4809(68)90018-9

Smestad, T. L. (2018). *Personality Matters! Improving the User Experience of Chatbot Interfaces - Personality Provides a Stable Pattern to Guide the Design and Behaviour of Conversational Agents*. Master's thesis, NTNU, Trondheim.

Smith, A., and Williams, K. D. (2004). R u there? Ostracism by cell phone text messages. *Group Dyn.* 8, 291–297. doi: 10.1037/1089-2699.8.4.291

Sproull, L., Subramani, M., Kiesler, S., Walker, J., and Waters, K. (1996). When the interface is a face. *Hum. Comput. Inter.* 11, 97–124. doi: 10.1207/s15327051hci1102_1

Stillman, T. F., Baumeister, R. F., Lambert, N. M., Crescioni, A. W., DeWall, C. N., and Fincham, F. D. (2009). Alone and without purpose: life loses meaning following social exclusion. *J. Exp. Soc. Psychol.* 45, 686–694. doi: 10.1016/J.JESP.2009.03.007

Sundar, S., and Nass, C. (2000). Source orientation in human-computer interaction. *Commun. Res.* 27, 683–703. doi: 10.1177/009365000027006001

Tantam, D. (2006). The machine as psychotherapist: impersonal communication with a machine. *Adv. Psychiatr. Treat.* 12, 416–426. doi: 10.1192/apt.12.6.416

Terada, K., Jing, L., and Yamada, S. (2015). "Effects of agent appearance on customer buying motivations on online shopping sites," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*, (New York, NY: ACM Press), 929–934. doi: 10.1145/2702613.2732798

Thompson, E. R. (2007). development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *J. Cross Cult. Psychol.* 38, 227–242. doi: 10.1177/0022022106297301

Tulin, M., Pollet, T., and Lehmann-Willenbrock, N. (2018). Perceived group cohesion versus actual social structure: a study using social network analysis of egocentric Facebook networks. *Soc. Sci. Res.* 74, 161–175. doi: 10.1016/j.ssresearch.2018.04.004

Tversky, A., and Kahneman, D. (1974). Judgement under uncertainty: heuristics and biases. *Science* 185, 1124–1131. doi: 10.1126/science.185.4157.1124

Ussher, J., Kirsten, L., Butow, P., and Sandoval, M. (2006). What do cancer support groups provide which other supportive relationships do not? The experience of peer support groups for people with cancer. *Soc. Sci. Med.* 62, 2565–2576. doi: 10.1016/J.SOCSCIMED.2005.10.034

Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., and Bickmore, T. (2012). "Designing relational agents as long term social companions for older adults," in *Proceedings of the International Conference on Intelligent Virtual Agents*, (Berlin: Springer), 289–302. doi: 10.1007/978-3-642-33197-8_30

Wada, K., and Shibata, T. (2007). Living with seal robots–its sociopsychological and physiological influences on the elderly at a care house. *IEEE Trans. Robot.* 23, 972–980. doi: 10.1109/tro.2007.906261

Wang, Y., Khooshabeh, P., and Gratch, J. (2013). "Looking real and making mistakes," in *Proceedings of International Workshop on Intelligent Virtual Agents*, Honolulu, HI, 339–348. doi: 10.1007/978-3-642-40415-3_30

Wanner, L., André, E., Blat, J., Dasiopoulou, S., Farrús, M., Fraga, T., et al. (2017). Design of a knowledge-based agent as a social companion. *Proc. Comput. Sci.* 121, 920–926.

Waytz, A., Cacioppo, J., and Epley, N. (2010). Who sees human? *Perspect. Psychol. Sci.* 5, 219–232. doi: 10.1177/1745691610369336

Weizenbaum, J. (1967). Contextual understanding by computers. *Commun. ACM* 10, 474–480. doi: 10.1145/363534.363545

Williams, K. D. (2001). *Ostracism: The Power of Silence*. New York, NY: Guilford Press.

Williams, K. D. (2007). Ostracism: the kiss of social death. *Soc. Pers. Psychol. Compass* 1, 236–247. doi: 10.1111/j.1751-9004.2007.00004.x

Williams, K. D. (2009). Ostracism: a temporal need-threat model. *Adv. Exp. Soc. Psychol.* 41, 275–314. doi: 10.1016/S0065-2601(08)00406-1

Williams, K. D., Cheung, C. K. T., and Choi, W. (2000). Cyberostracism: effects of being ignored over the internet. *J. Pers. Soc. Psychol.* 79, 748–762. doi: 10.1037/0022-3514.79.5.748

Williams, K. D., Forgas, J. P., Hippel, W., Von, Forgas, J. P., Hippel, W., et al. (2005). *The Social Outcast*. London: Psychology Press.

Wolf, W., Levordashka, A., Ruff, J. R., Kraaijeveld, S., Lueckmann, J.-M., and Williams, K. D. (2015). Ostracism online: a social media ostracism paradigm. *Behav. Res. Methods* 47, 361–373. doi: 10.3758/s13428-014-0475-x

Zadro, L., Williams, K. D., and Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *J. Exp. Soc. Psychol.* 40, 560–567. doi: 10.1016/J.JESP.2003.11.006

Zamora, J. (2017). "I'm sorry, Dave, I'm afraid I can't do that," in *Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17*, (New York, NY: ACM Press), 253–260. doi: 10.1145/3125739.3125766

Zhang, X., Norris, S. L., Gregg, E. W., and Beckles, G. (2007). Social support and mortality among older persons with diabetes. *Diabetes Educ.* 33, 273–281. doi: 10.1177/0145721707299265

Zhong, C.-B., and Leonardelli, G. J. (2008). Cold and lonely. *Psychol. Sci.* 19, 838–842. doi: 10.1111/j.1467-9280.2008.02165.x

Zumstein, D., and Hundertmark, S. (2017). Chatbots: an interactive technology for personalized communication, transactions and services. *IADIS Int. J. Internet* 15, 96–109.