



Effect of AI chatbot emotional disclosure on user satisfaction and reuse intention for mental health counseling: a serial mediation model

Gain Park¹ · Jiyun Chung² · Seyoung Lee³

Accepted: 25 October 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

This study explored the effect of chatbot emotional disclosure on user satisfaction and reuse intention for a chatbot counseling service. It also examined the independent and sequential mediation roles of user emotional disclosure intention and perceived intimacy with a chatbot on the relationship between chatbot emotional disclosure, user satisfaction, and reuse intention for chatbot counseling. In total, 348 American adults were recruited to participate in a mental health counseling session with either of the two types of artificial intelligence-powered mental health counseling chatbots. These included a chatbot disclosing factual information only or a chatbot disclosing humanlike emotions. The results revealed that chatbot emotional disclosure significantly increased user satisfaction and reuse intention for a chatbot counseling service. The results further revealed that user emotional disclosure intention and perceived intimacy with a chatbot independently and serially mediate the effect of chatbot emotional disclosure on user satisfaction and chatbot counseling service reuse intention. The results indicate positive effects of artificial emotions and their disclosure in the context of chatbot moderated mental health counseling. Practical implications and psychological mechanisms are discussed.

Keywords Chatbot · Emotional disclosure · Mental health · Intimacy · User satisfaction · Intention to reuse

The labor shortage has long been a challenge in the mental health services industry. One solution has been to adopt artificial intelligence (AI) chatbots for mental health counseling services. Mental health counseling chatbots such as Woebot and Gabby have proven to reduce symptoms of depression

and automate some parts of clinical treatment, showing promising results for the development of AI chatbot counselors (Fitzpatrick et al., 2017; Miner et al., 2017). The professionals also agree that chatbot counselors can help people improve their mental health in the early stage, however, there still exist concerns regarding chatbots' ability to display a range of human emotions in sensitive conversational situations and encourage emotional disclosure of counselees, which are essential parts of mental health counseling (Burch, 2004; Reis & Franks, 1994; Sweeney et al., 2021).

According to social penetration theory, one's emotional disclosure is positively associated with reciprocal emotional disclosure and perceived intimacy with a conversational partner (Carpenter & Greene, 2016). Based on the computers are social actors (CASA) paradigm, this study argues that social rules in human–human interactions can be applied to human–chatbot interactions (Nass & Moon, 2000; Reeves & Nass, 1996), suggesting that the norm of reciprocity would also be valid for artificially generated emotional disclosures.

However, another line of studies has suggested possible negative effects of artificial human likeness due to

✉ Seyoung Lee
gethemane@skku.edu

Gain Park
park@nmsu.edu

Jiyun Chung
chungjiy@gmail.com

¹ Department of Journalism and Media Studies, New Mexico State University, 2915 McFie Circle, Milton Hall 158, Las Cruces, NM 88003, USA

² Convergence and Open Sharing System-Artificial Intelligence, Sungkyunkwan University, 25-2 Sungkyunkwan-Ro, 50212 Hoam Hall, Jongno-Gu, Seoul, South Korea 03063

³ Department of Media and Communication, Sungkyunkwan University, 25-2, Sungkyunkwan-Ro, 50505 Hoam Hall, Jongno-Gu, Seoul, South Korea 03063

uncanniness or perceived deception (e.g., Liu & Sundar, 2018; Mori et al., 2012; Skjuve et al., 2019). We believe such inconsistencies in the discussion of artificial human likeness and a lack of studies on chatbot emotions call for conducting further research. To that end, we developed two AI-powered chatbot counselors (factual vs. emotional disclosure) and conducted a quasi-experiment. This study provides theoretical and practical implications for human–computer interaction researchers and practitioners in the mental health services industry.

Theoretical premises

Effect of chatbot emotional disclosure on user satisfaction and reuse intention

Emotional disclosure is critical to interpersonal communication and is an important strategy in mental health counseling. Effective and timely emotional responses are key to increasing user satisfaction and positively impacting mental health and offer mental health counselors the ability to exchange feelings and meanings with a client (Clark, 2010). Moreover, for counselees with mental health issues, failing to provide emotional support in the early stage can significantly reduce user satisfaction and future help-seeking behavior (Miner et al., 2017). Social penetration theory further suggests several positive outcomes of emotional disclosure such as increased closeness, trust, enjoyment, and intimacy with a conversational partner in various contexts (Carpenter & Greene, 2016; Ho et al., 2018; Lee et al., 2020). Malloch and Zhang (2019) also suggested the effect of self-disclosure on health behavioral intention; they tested the effects of varying levels of self-disclosure in posters from an online support group and found that cognitive and emotional disclosure led to higher perceived support availability and dieting intention.

Studies suggest that the same effect can be extended to human–computer interaction (Brandtzaeg & Følstad, 2017; Go & Sundar, 2019; Ho et al., 2018; Moon, 2000; Nass & Moon, 2000; Reeves & Nass, 1996). The CASA paradigm (Reeves & Nass, 1996) and the social response theory (Moon, 2000) point out that people automatically apply social rules and norms while interacting with computers. Nass and Moon (2000) also suggest that people mindlessly apply social scripts that they use in human-to-human interaction, ignoring other nonhuman cues of computer partners. Moreover, extant studies posit that people are motivated not only by productive reasons but also by social motivations, in that people are motivated by friendly and empathic relationships with chatbots (Brandtzaeg & Følstad, 2017; Choi & Drumwright, 2021), suggesting the positive effect of emotional disclosures in human-chatbot interaction. In line

with the discussion, studies revealed that several social cues can help activate humanlike interaction and yield positive behavioral outcomes (Brandtzaeg & Følstad, 2017; Go & Sundar, 2019; Ho et al., 2018; Moon, 2000; Nass & Moon, 2000; Reeves & Nass, 1996), which can be categorized into conversational, visual, auditory, and invisible aspects (Feine et al., 2019). For example, Go and Sundar (2019) found that simple changes in conversational styles of chatbots increase social presence, perceived similarity, and other attitudinal and behavioral outcomes. Moreover, Ho and colleagues (Ho et al., 2018) found that emotional disclosure had positive emotional, relational, and psychological outcomes in the context of human-chatbot conversations; the effects of emotional disclosure were equivalent regardless of whether people perceived their conversational partner as a human or a chatbot.

Based on the above, this study aimed to examine the effectiveness of emotional disclosure in the context of chatbot mental health counseling and investigate whether disclosing computer-generated humanlike emotions can have the same effect as human emotional disclosure. To that end, we adopted the understanding of self-disclosure as being in two streams (i.e., factual vs. emotional disclosure) and proposed that chatbots' emotional disclosure will yield higher user satisfaction and reuse intention for a chatbot counseling service.

H1: A chatbot with emotional disclosure will yield higher user satisfaction and reuse intention for a chatbot counseling service than a chatbot with factual disclosure.

Mediating the role of user emotional disclosure intention

Another important implication of social penetration theory is the norm of reciprocity. The theory suggests that self-disclosure can make conversational partners feel obliged to keep the norm of equity and match up the level of disclosure to their conversational partners' disclosure level (Carpenter & Greene, 2016). For example, if one shares his or her mental health struggles, such as feelings of sorrow, anxiety, and helplessness with another person, the receivers of such information will be more likely to share their feelings and experience. Lee and colleagues (Lee et al., 2020) found that such reciprocity can too occur in human-chatbot conversations. They created three chatbots with no disclosure, low disclosure, and high disclosure conditions, asked participants to engage in a conversation with one of the chatbots, and asked some sensitive questions at the end of the conversation to gauge if the chatbot's level of disclosure affected users' openness toward the questions (Lee et al., 2020). The results provided evidence of reciprocity; participants were more likely to share sensitive information (e.g., "I was emotionally

abused by my ex-boyfriend”) after being exposed to a higher level of chatbot self-disclosure (Lee et al., 2020), which then led to conversational enjoyment.

Uncertainty reduction theory further explains that such emotional exchanges reduce uncertainty in an early-stage relationship and increase the level of emotional exchange (Berger & Calabrese, 1975). This has important implications for the mental health context as one of the key purposes of early-stage mental health counseling is to promote users’ emotional disclosure by reducing the level of risk from uncertainty (see also the revelation risk model; Afifi & Steuber, 2009). Moreover, such emotional exchanges can create a feeling of being understood, which may result in successful psychotherapeutic alliances (Elliott et al., 2011), satisfaction and behavior change (Sanford, 2006), improved mental and physical health (Lepore et al., 2004), and other benefits (Reis et al., 2017).

According to the CASA paradigm (Reeves & Nass, 1996) and the social response theory (Moon, 2000), we posit that the above-explained mechanism between chatbot emotional disclosure, user emotional disclosure, and outcome variables can be applied to human-chatbot interaction. However, we argue that there is still limited evidence to support its application in the context of human-chatbot mental health counseling, and not many studies have examined its practical outcomes such as user satisfaction and reuse intention for a chatbot counseling service. Therefore, our second hypothesis proposes that user emotional disclosure intention will mediate the effect of chatbot emotional disclosure on user satisfaction and reuse intention for a chatbot counseling service.

H2: User emotional disclosure intention will mediate the effect of chatbot emotional disclosure on user satisfaction and reuse intention for a chatbot counseling service.

Mediating the role of perceived intimacy

Intimacy has been defined as “the ambiance of closeness” or “shared experience of openness” (Levenson, 1974, pp. 359, 368). Extant research has found a positive effect of emotional disclosure on perceived intimacy with a conversational partner (Carpenter & Greene, 2016; Ho et al., 2018; Lee et al., 2020), which then leads to several positive outcomes such as lower depression (Reis & Franks, 1994), and higher social support satisfaction (Johnson et al., 1993), perceived understanding (Reis et al., 2017), enjoyment of self-insight and personal growth (Reis & Shaver, 1988), purchase intention (Yin et al., 2019), continuous intention (Lin et al., 2021), and repurchase intention (Huaman-Ramirez et al., 2022).

Studies have also found positive associations between self-disclosure, intimacy, and several beneficial outcomes in the context of chatbot and computer-mediated

communications (e.g., Ho et al., 2018; Lee et al., 2020; Moon, 2000). For example, Sung and colleagues (Sung et al., 2007) found that developing intimacy with a vacuuming robot increased the pleasure of cleaning, efforts to fit the robot into their homes, and recommendation intentions. The findings of human–computer interaction studies that people are motivated by friendly and empathic relationships with computer partners also support the effect of perceived intimacy on outcome variables (Brandtzaeg & Følstad, 2017; Choi & Drumwright, 2021). Brandtzaeg and Følstad (2017) point out that social or relational motivations can significantly improve user experience, and people not only seek productivity from chatbots but desire more intimate relationships in human-chatbot interaction, suggesting a positive relationship between perceived intimacy and interaction outcomes.

We argue that this psychological mechanism can be applied to the role of perceived intimacy with a chatbot counselor in enhancing user satisfaction and reuse intention for a chatbot counseling service. Therefore, we propose the following hypothesis.

H3: Perceived intimacy with a chatbot will mediate the effect of chatbot emotional disclosure on user satisfaction and reuse intention for a chatbot counseling service.

Serial mediation through user emotional disclosure intention and perceived intimacy

In the previous sections, we have discussed the effect of chatbot emotional disclosure on user emotional disclosure, the effect of user emotional disclosure intention on perceived intimacy, and the effect of perceived intimacy on user satisfaction and reuse intention. This study further proposes a serial mediation between the variables such that user emotional disclosure intention and perceived intimacy with a chatbot serially mediate the relationship between chatbot emotional disclosure and the dependent variables (i.e., user satisfaction and reuse intention). Extant studies have implied the serial mediation of the reciprocity of self-disclosure (Carpenter & Greene, 2016; Lee et al., 2020) and pointed out that in therapy, intimacy requires patients’ risk-taking with regard to their private information and helps build authentic relationships (Levenson, 1974), suggesting the serial order of the mediation effects.

To further the understanding of the psychological mechanism between chatbot emotional disclosure and the outcome variables, we propose the following hypothesis to examine the serial mediation of user emotional disclosure intention and perceived intimacy with a chatbot counselor on the relationship between chatbot emotional disclosure and user satisfaction/reuse intention. We employed self-reported user emotional disclosure intention instead of actual user

emotional disclosure. The rationale for this will be discussed in the discussion section.

H4: User emotional disclosure intention and intimacy with a chatbot will serially mediate the effect of chatbot emotional disclosure on user satisfaction and reuse intention for a chatbot counseling service.

Method

Participants

A total of 360 American adults were recruited in March 2022 through Amazon Mechanical Turk, a crowdsourcing website. All participation was voluntary and informed consent forms were signed before data collection. Participants who did not complete the conversation with the chatbot or the survey questionnaire were excluded from the analysis, yielding a final sample of 348 participants. In total, 177 (50.9%) participants were assigned to the factual disclosure only condition and 171 (49.1%) participants were assigned to the emotional disclosure condition. The total sample consisted of 190 (54.6%) male and 158 (45.4%) female participants; comprising 234 White (67.2%), 40 African American (11.5%), 48 Asian (13.8%), and 24 Native Indian/Alaska Native (6.9%) participants, and 2 participants from other races (0.6%). The participants' average age was 36.9 (min = 21, max = 71). Seventy participants had completed high school and 278 held a bachelor's degree or higher. The median household annual income ranged from \$50,000 to \$59,999.

Stimulus

Two types of chatbots (i.e., factual vs. emotional disclosure) were developed based on the design of Woebot, an AI-powered mental health chatbot counselor (Mack, 2018). In the factual condition, the chatbot counselor only provided the factual information without disclosing any emotional information (e.g., "Experts recommend connecting with others in supportive conversation is an important self-care practice"). The factual information was designed based on the Centers for Disease Control and Prevention's (CDC, n.d.) mental health recommendations and modified under the close supervision of a certified mental health practitioner. In contrast, in the emotional disclosure condition, the chatbot counselor disclosed its own emotional experience (e.g., "I felt much better after connecting with others in supportive conversation"). The emotional disclosures were designed based on Malloch and Zhang's (2019) and Ho and colleagues' (Ho et al., 2018) studies to illustrate the chatbot's feelings and emotional reflections.

After participants completed the pre-questionnaire, they were asked to click a link taking them to a chatbot counseling page with either the factual information or emotional disclosure conditions. For both the factual information and emotional disclosure conditions, participants were greeted by a chatbot counselor ("Hello, I'm your virtual mental health counselor. I'm here to discuss your mental health. Before we begin, can you answer some questions about your mental health conditions?") and asked to complete a simple mental health questionnaire developed based on the CDC's mental health guidelines (CDC, n.d.; e.g., "How well could you enjoy your favorite activities?"). Throughout the counseling session, participants used button-type options to respond to the chatbot's questions and statements (e.g., "I'm doing quite well" or "I'm not doing very well"), thus employing the conversational design of Woebot (Fitzpatrick et al., 2017) to avoid technical issues or errors caused by typos or unexpected answers (see Fig. 1). Participants were then provided with mental health information and suggestions based on the CDC's mental health guidelines (CDC, n.d.). The entire conversation between the chatbot and participants was developed under the supervision of a certified mental health practitioner. In both conditions, participants were thanked and debriefed with a verification code and directed back to the survey page to complete the post-questionnaire.

Procedure

Two different chatbots were developed using Dialogflow, an automated text-based program with a natural language processing engine. This program allowed AI chatbots and users to interact in a two-way conversation throughout the counseling session by generating appropriate responses to users' inputs. The two chatbots were trained either to practice factual or emotional disclosure during counseling while they ask basic mental health-related questions and make mental health suggestions to participants. The mental health-related questions and suggestions were developed using the CDC's (n.d.) mental health resources and reviewed by a certified counselor. An online survey questionnaire containing a link to either version of the chatbot counselor was constructed using REDCap software and distributed through Amazon Mechanical Turk.

The first part of the questionnaire contained a pre-questionnaire with an informed consent form and included questions about participants' previous mental health-related experiences. Participants were then asked to read a scenario in which they are going to have a counseling session with a chatbot to address mental health problems amid COVID-19. After reading the scenario, participants were randomly assigned to either condition of the chatbot (factual vs. emotional disclosure) and asked to click a link

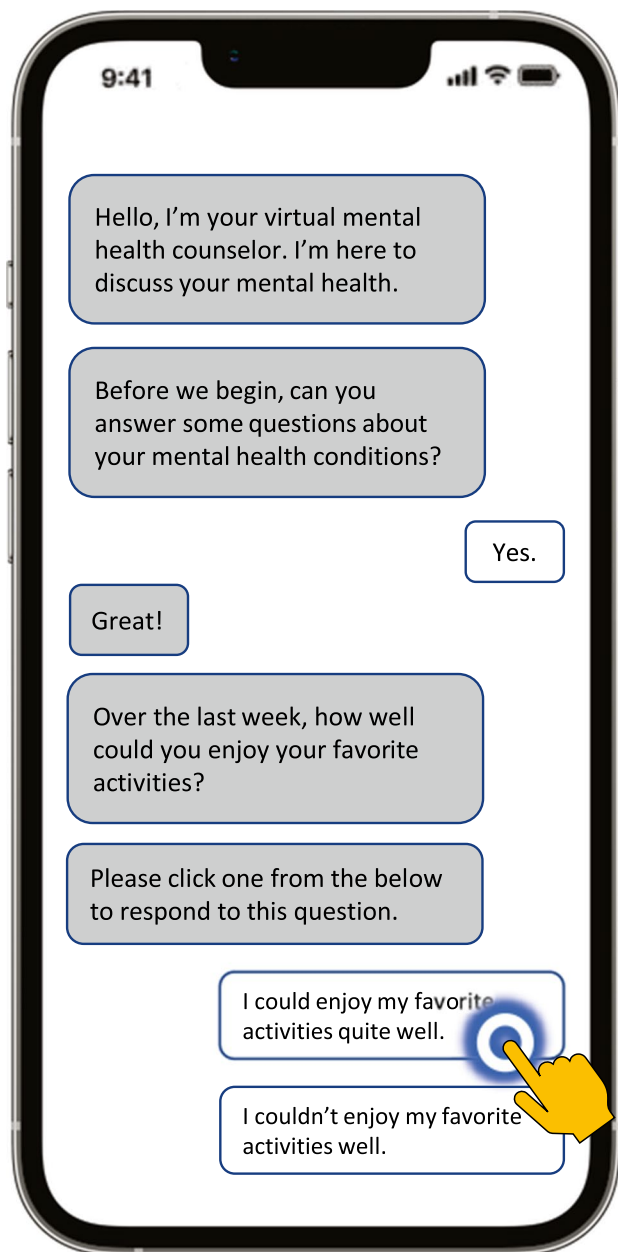


Fig. 1 Button-based chatbot screen

leading them to an online chat page. At the end of the counseling session, participants were provided a verification code as proof of completion.

After a counseling session with a chatbot, participants were led back to the questionnaire page and answered the post-questionnaire that included items on perceived chatbot emotional disclosure (manipulation check), user emotional disclosure intention, perceived intimacy with the chatbot, user satisfaction, and reuse intention. Next, participants were asked questions related to their demographics and then debriefed.

Measurements

Chatbot emotional disclosure (manipulation check)

An emotional disclosure scale was adopted from Ho and colleagues' study (Ho et al., 2018) and modified to check whether the two chatbot conditions (factual vs. emotional disclosure) were manipulated as intended. For example, participants reported how much the counselor "expressed personal emotion" and "shared its feelings" during the counseling session. Each item was rated on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree ($M = 4.67$, $SD = 1.43$, Cronbach's $\alpha = 0.93$).

User emotional disclosure intention

User emotional disclosure intention was evaluated through five self-reported items including "I was willing to discuss my feelings with the counselor" and "If I have another counseling session with this counselor, I would be willing to share my feelings" (Ho et al., 2018; Lee & Choi, 2017; Malloch & Zhang, 2019). Each item was rated on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree ($M = 5.01$, $SD = 1.04$, Cronbach's $\alpha = 0.88$).

Perceived intimacy with a chatbot

We adopted Levenson's (1974) view and defined intimacy as the perceived closeness between a counselor and a patient and modified the intimacy scale developed by Lee and Choi (2017) to measure intimacy with the chatbot counselor (e.g., "I felt close to the counselor during the conversation" and "I experienced the interaction as intimate"). Each item was rated on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree ($M = 4.37$, $SD = 1.56$, Cronbach's $\alpha = 0.93$).

User satisfaction with chatbot counseling

User satisfaction with chatbot counseling was assessed using six questions including "I am satisfied with the counselor's recommendation service" and "Interacting with the counselor was a pleasant and satisfactory experience" (Lee & Choi, 2017). Each item was rated on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree ($M = 4.99$, $SD = 1.08$, Cronbach's $\alpha = 0.90$).

Reuse intention for chatbot counseling

Reuse intention for chatbot counseling was measured using four questions including "I am willing to use the counseling service again" and "I intend to speak with the counselor again when I feel depressed" (Lee & Choi, 2017). Each item

was rated on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree ($M = 4.79$, $SD = 1.33$, Cronbach's $\alpha = 0.91$).

Statistical analyses

We first conducted descriptive statistics, Cronbach's α , and correlation analysis using SPSS 26.0. Next, we conducted independent samples t -tests to examine the difference in the two types of chatbots on user satisfaction and reuse intention for chatbot counseling (H1). Then, to test the mediation roles of user emotional disclosure intention and user intimacy with a chatbot counselor on the relationships between chatbot emotional disclosure and user satisfaction and reuse intention for chatbot counseling (H2, H3, and H4), we conducted the PROCESS macro Model 6 (Hayes, 2013). The bootstrapping method condition generated 95% biased-corrected confidence intervals (CI) based on 5,000 bootstrap samples.

Results

Manipulation check

Participants in the emotional disclosure condition ($M = 5.37$, $SD = 0.93$) perceived that the chatbot counselor revealed more emotion to them than did those in the factual information condition ($M = 3.99$, $SD = 1.49$), $t(346) = 10.24$, $p < 0.001$. Therefore, the two chatbot conditions (factual information vs. emotional disclosure) were manipulated as intended.

Preliminary analyses

Table 1 presents the descriptive statistics and zero-order correlations for all the variables considered in this study. All correlation coefficients among the variables were statistically significant.

Hypothesis testing

H1 posited that the chatbot counselor that discloses its own emotional experience in the mental health context would yield higher user satisfaction and reuse intention for the chatbot counseling service than one that only provides factual information without disclosing any emotional information. To test H1, we performed an independent samples t -test. The results demonstrated that participants in the emotional disclosure condition ($M = 5.41$, $SD = 0.74$) reported higher user satisfaction with the chatbot counseling service than those in the factual information condition ($M = 4.59$, $SD = 1.20$), $t(346) = 7.58$, $p < 0.001$. Additionally, participants in the emotional disclosure condition ($M = 5.34$, $SD = 0.92$) indicated higher intention to reuse the chatbot counselor than those in the factual information condition ($M = 4.26$, $SD = 1.44$), $t(346) = 8.24$, $p < 0.001$. Therefore, H1 was supported.

We used Hayes' (2013) PROCESS macro Model 6 to test H2, H3, and H4. The chatbot emotional disclosure (0 = factual information, 1 = emotional disclosure) was entered as an independent variable, user emotional disclosure intention and user intimacy with a chatbot counselor were entered as mediators, and user satisfaction and reuse intention for the chatbot counseling service were entered as dependent variables.

H2 postulated that user emotional disclosure intention would mediate the relationship between chatbot emotional disclosure and user satisfaction and reuse intention for the chatbot counseling service. As shown in Fig. 1 and Table 2, chatbot emotional disclosure was positively associated with user emotional disclosure intention ($B = 0.72$, $SE = 0.11$, 95% $CI = 0.51, 0.92$), which in turn was positively related to user satisfaction with the chatbot counseling service ($B = 0.52$, $SE = 0.05$, 95% $CI = 0.43, 0.60$) and user intention to reuse the chatbot counseling service ($B = 0.47$, $SE = 0.05$, 95% $CI = 0.38, 0.56$). Thus, the impact of chatbot emotional disclosure on user satisfaction with the chatbot counseling service ($B = 0.37$, $Boot SE = 0.08$, 95% $Boot CI = 0.23, 0.54$) and user intention to reuse the chatbot counseling service

Table 1 Descriptive statistics for and zero-order correlations between all variables in the model

	M (SD)	1	2	3	4	5
1. CED	-	1.00				
2. UEDI	5.00 (1.04)	.34**	1.00			
3. PIC	4.37 (1.56)	.37**	.59**	1.00		
4. USCC	4.99 (1.08)	.38**	.71**	.64**	1.00	
5. IRCC	4.79 (1.33)	.41**	.72**	.79**	.82**	1.00

** $p < .01$

CED=Chatbot Emotional Disclosure, UEDI=User Emotional Disclosure Intention, PIC=Perceived Intimacy with Chatbot, USCC=User Satisfaction with Chatbot Counseling, IRCC=Intention to Reuse Chatbot Counseling

Table 2 Total, direct, and indirect effects of CED on USCC/IRCC

USCC outcomes as criterion					
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>LLCI</i>	<i>ULCI</i>
Total effect of CED on USCC	.81	.11	7.58	.60	1.03
Direct effect of CED on USCC	.20	.08	2.43	.04	.36
Indirect effects of CED on USCC					
	<i>B</i>	<i>Boot SE</i>		<i>Boot LLCI</i>	<i>Boot ULCI</i>
Total indirect effect of CED on USCC	.61	.09		.44	.80
Indirect effect 1: CED → UEDI → USCC	.37	.08	-	.22	.54
Indirect effect 2: CED → PIC → USCC	.12	.05	-	.05	.23
Indirect effect 3: CED → UEDI → PIC → USCC	.12	.03	-	.06	.19
IRCC outcomes as criterion					
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>LLCI</i>	<i>ULCI</i>
Total effect of CED on IRCC	1.08	.13	8.24	.82	1.33
Direct effect of CED on IRCC	.21	.08	2.54	.05	.37
Indirect effects of CED on IRCC					
	<i>B</i>	<i>Boot SE</i>		<i>Boot LLCI</i>	<i>Boot ULCI</i>
Total indirect effect of CED on IRCC	.87	.12		.64	1.10
Indirect effect 1: CED → UEDI → IRCC	.34	.08	-	.21	.50
Indirect effect 2: CED → PIC → IRCC	.27	.08	-	.12	.42
Indirect effect 3: CED → UEDI → PIC → IRCC	.26	.05	-	.17	.37

CED = Chatbot Emotional Disclosure, UEDI = User Emotional Disclosure Intention, PIC = Perceived Intimacy with Chatbot, USCC = User Satisfaction with Chatbot Counseling, IRCC = Intention to Reuse Chatbot Counseling. The number of bootstrap samples for bias-corrected bootstrap confidence intervals: 5,000. The level of confidence for all confidence intervals: 95

($B = 0.34$, $Boot SE = 0.07$, 95% $Boot CI = 0.20, 0.50$) was mediated by user emotional disclosure. Specifically, the chatbot counselor who disclosed its own emotions resulted in higher user emotional disclosure, which then yielded higher user satisfaction and reuse intention. Thus, H2 was supported.

H3 posited that user intimacy with the chatbot counselor would mediate the relationship between chatbot emotional disclosure and user satisfaction and reuse intention for the chatbot counseling. The findings indicated that chatbot emotional disclosure was positively associated with user intimacy with the chatbot counselor ($B = 0.58$, $SE = 0.14$, 95% $CI = 0.30, 0.86$), which in turn was positively related to user satisfaction with the chatbot counseling service ($B = 0.22$, $SE = 0.03$, 95% $CI = 0.16, 0.28$) and user intention to reuse the chatbot counseling service ($B = 0.46$, $SE = 0.03$, 95% $CI = 0.40, 0.52$). Therefore, the influence of chatbot emotional disclosure on user satisfaction with the chatbot counseling ($B = 0.13$, $Boot SE = 0.05$, 95% $Boot CI = 0.05, 0.23$) and user intention to reuse the chatbot counseling service ($B = 0.27$, $Boot SE = 0.08$, 95% $Boot CI = 0.12, 0.43$) was mediated by user intimacy with the chatbot counselor. That is, the chatbot counselor who disclosed its own emotions generated higher user intimacy, which then led to higher user satisfaction and intention to reuse. Hence, H3 was supported.

H4 suggested that user emotional disclosure intention and user intimacy with the chatbot counselor would serially mediate the relationship between chatbot emotional disclosure and user satisfaction and user intention to reuse the

chatbot counseling service. The results showed that chatbot emotional disclosure was positively related to user emotional disclosure intention ($B = 0.72$, $SE = 0.11$, 95% $CI = 0.51, 0.92$), user emotional disclosure intention was positively related to user intimacy with the chatbot counselor ($B = 0.79$, $SE = 0.07$, 95% $CI = 0.66, 0.93$), and in turn, user intimacy with the chatbot counselor was positively related to user satisfaction with the chatbot counseling ($B = 0.22$, $SE = 0.03$, 95% $CI = 0.16, 0.28$) and intention to reuse the chatbot counseling service ($B = 0.46$, $SE = 0.03$, 95% $CI = 0.40, 0.52$). That is, when the chatbot counselor disclosed its own emotions, it led to higher user emotional disclosure, which in turn created higher user intimacy with the chatbot counselor. This then led to user satisfaction with the chatbot counseling ($B = 0.12$, $Boot SE = 0.03$, 95% $Boot CI = 0.06, 0.19$) and user intention to reuse the chatbot counseling ($B = 0.26$, $Boot SE = 0.05$, 95% $Boot CI = 0.17, 0.37$). Therefore, H4 was supported (see Fig. 2).

Discussion

The results suggest that chatbot emotional disclosure is positively associated with user satisfaction and reuse intention for chatbot counseling services in which people who are exposed to chatbot emotional disclosure would show higher user satisfaction and reuse intention than those who are only exposed to factual information. We also found that user emotional disclosure intention and perceived intimacy with a

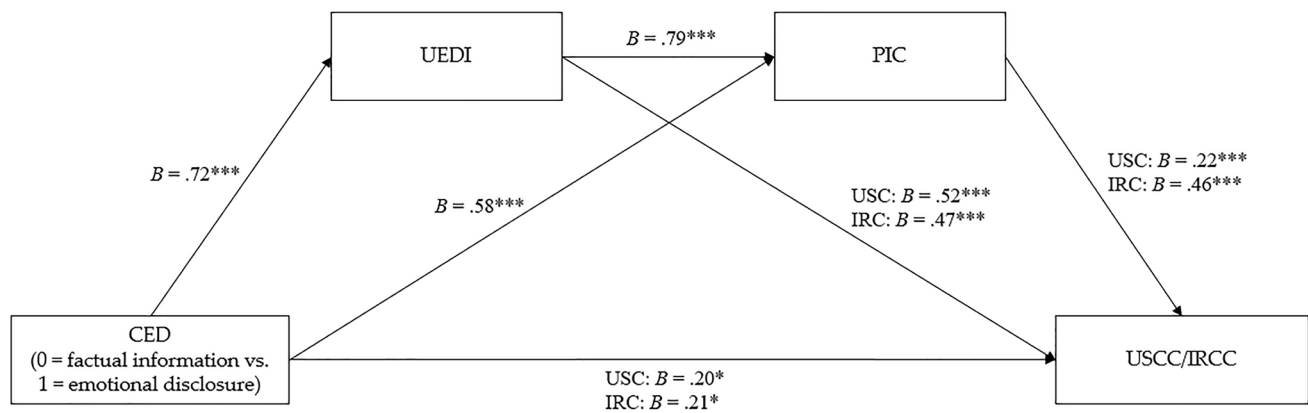


Fig. 2 Unstandardized path coefficients for serial multiple mediation. CED=Chatbot Emotional Disclosure, UEDI=User Emotional Disclosure Intention, PIC=Perceived Intimacy with Chatbot,

USCC=User Satisfaction with Chatbot Counseling, IRCC=Intention to Reuse Chatbot Counseling * $P < .05$, *** $p < .001$

chatbot would mediate the relationship between chatbot emotional disclosure and user satisfaction and reuse intention. Moreover, the results further suggest that user emotional disclosure intention and perceived intimacy serially mediate the effect of chatbot emotional disclosure on user satisfaction and reuse intention.

These findings support the notion of the CASA framework that people assign humanlike characteristics to computers and technology and abide by social rules as they do in human relationships (Reeves & Nass, 1996). The major contribution of this paper is that it has clarified the psychological mechanism underlying the CASA framework by examining the serial mediation of user emotional disclosure and perceived intimacy. Social penetration theory supports our findings that emotional disclosure compels a conversational partner to have the same level of disclosure to maintain reciprocity (Carpenter & Greene, 2016), which then leads to higher intimacy with the conversational partner (Carpenter & Greene, 2016; Fehr, 2004; Lee et al., 2020) and various beneficial outcomes such as lower depression (Reis & Franks, 1994) and higher social support satisfaction (Johnson et al., 1993), perceived understanding (Reis et al., 2017), enjoyment of self-insight and personal growth (Reis & Shaver, 1988), purchase intention (Yin et al., 2019), continuous intention (Lin et al., 2021), and repurchase intention (Huaman-Ramirez et al., 2022). Uncertainty reduction theory further supports the effect of initial emotional disclosure in reducing uncertainty in an early-stage relationship and increasing the level of emotional exchange (Berger & Calabrese, 1975), which can lead to a feeling of being understood and successful psychotherapeutic alliances (Elliott et al., 2011), satisfaction and behavior change (Sanford, 2006), and other benefits (Reis et al., 2017). We argue that revealing the psychological process of the perception and acceptance of humanlike chatbot designs is the key to clarifying the

reasons for the conflicting results in robot humanness studies. Moreover, to the best of our knowledge, this is the first attempt to test the serial mediation through user emotional disclosure intention and perceived intimacy in the context of chatbot conversation. We believe this research can contribute to the understanding of human–computer integration, especially in the context of chatbot counseling.

Another major contribution of this paper is that it expands the scope of the theoretical framework on emotional disclosures and exchanges in the context of chatbot conversations and artificial emotions. Consistent with the CASA framework, we found that in the context of mental health counseling, artificially created emotions function similarly to human emotions and yield various conversational benefits such as increased willingness to share emotions, perceived intimacy, user satisfaction, and reuse intention. This result is in contrast to the assumption that chatbots lack in performing emotional tasks and that humanlike chatbots may create uncanniness or discomfort (Ciechanowski et al., 2019). We believe such conflicting results may have been caused by contextual differences; that is in the context of mental health counseling, people tend to expect more emotional exchanges with chatbot agents than in other contexts such as consumer services. We argue that people mindlessly apply higher social expectations to chatbot counselors as they would do in human-to-human counseling sessions, ignoring the known fact that the counselor is an AI chatbot agent who is not capable of feeling or having emotions (Moon, 2000; Nass & Moon, 2000). We suggest future studies to test the hypotheses and compare the results in different contexts with varying levels of expected emotional disclosures.

We also believe that the findings have important practical implications for stakeholders in the field of mental health counseling, where there are active attempts to replace human counselors with AI chatbots and utilize artificial emotions to

provide emotional support for counselees due to labor shortages and increased demand for mental health counselors. We also argue that the findings have important implications, especially in the field of mental health counseling, where patient disclosure, intimacy building, patient satisfaction, and revisit intention are considered major factors for successful counseling. Even though emotional disclosures and intimacy building are important factors in counseling, it is not always easy to earn them because the process includes patients' risk-taking with their private information (Levenson, 1974). This study's findings can help chatbot developers and mental health experts to build authentic relationships with counselees using chatbot counselors. Moreover, we believe that understanding how to increase openness and intimacy from chatbot mental health counseling would be especially helpful at this difficult time when people are mentally vulnerable and isolated due to the pandemic.

Despite this study's theoretical and practical implications, some limitations exist. First, during the chatbot counseling session, we limited users' responses to no text-based responses by providing clickable response options instead (e.g., "Quite well," "Not very well," "Yes, I have talked to a counselor," "No, I haven't talked to a counselor."). By employing this type of conversational design, we were able to minimize technical issues or errors caused by typos or unexpected answers while still giving participants some sense of being part of a conversation. Even though the design is an effective way of maintaining the intended flow of the conversation, it is limited in gauging emotional disclosure and the various expressions of counselees. This limitation led to the second issue, where we measured self-reported user emotional disclosure intention instead of actual disclosure that occurred during the conversation. This measure of disclosure has been employed by several studies (e.g., Chin et al., 2018; Lee & Choi, 2017; Wang et al., 2017), however, we believe future studies will benefit by qualitatively measuring users' actual emotional disclosure that occurs during counseling.

Another future direction would be to include data from other countries. Cultural differences and collectivistic and individualistic values can be important factors in shaping preference for emotional support (Burleson & Mortenson, 2003). Moreover, considering cultural contexts in the application of emotional disclosure is also a key factor in enhancing mental health, in that emotional disclosure can be perceived differently due to cultural variations (Lepore et al., 2004). There may also exist differences in the level of disclosure motivation due to the cultural context; people with collectivistic values tend to disclose less than those with individualistic values (Ting-Toomey, 1991). Hence, we believe future studies would benefit from expanding the data to include other countries and conducting cultural comparisons. We also suggest including gender and/or

age as variables in future studies. Studies pointed out that the gender stereotype of computers is a powerful cue and people apply the same gender stereotypes used in human-to-human interactions when interacting with computers (Nass & Moon, 2000; Nass et al., 1994). We further suggest that future studies should consider the effect of gender matching or mismatching on the therapeutic relationship. Another important demographic variable, age, should also be included in future studies as children are more likely to get emotionally involved with humanlike agents and be affected by them than adults (Vollmer et al., 2018).

Finally, we also suggest that future studies test the hypotheses in different social contexts beyond mental health counseling. Despite the evidence favoring the effect of emotional disclosure and the CASA framework (Reeves & Nass, 1996), studies also suggest that artificial human likeness should be used with caution as they may be perceived as deceptive or uncanny, especially for people who do not believe in the affective capacity of robots (Liu & Sundar, 2018; Mori et al., 2012; Portela & Granell-Canut, 2017; Skjerve et al., 2019). It will be beneficial to explore the conditions in which such an uncanny valley effect occurs. Testing the effect of chatbot emotional disclosure in a different context (e.g., supportive chatbot vs. persuasive chatbot) would enhance the understanding of the uncanny valley effect.

Conclusion

Based on social penetration theory, uncertainty reduction theory, and the CASA framework, an independent samples *t*-test found a significant positive effect of chatbot emotional disclosure on user satisfaction and intention to reuse a chatbot counseling service. Hayes' (2013) PROCESS macro Model 6 further revealed that both user emotional disclosure intention and perceived intimacy with a chatbot counselor mediate the relationship between chatbot emotional disclosure and user satisfaction and reuse intention for the chatbot counseling service. Moreover, we also found a serial mediation of user emotional disclosure intention and perceived intimacy with a chatbot counselor in the relationship between chatbot emotional disclosure and user satisfaction and reuse intention for the chatbot counseling service.

To summarize, the results suggest that chatbot emotional disclosure can enhance the effectiveness of chatbot-based mental health counseling by increasing user emotional disclosure and perceived intimacy, which are important aspects of mental health counseling. We expect the results to be helpful for chatbot developers and contribute to the mental health counseling industry that has suffered in the wake of COVID-19 and labor shortage. We also suggest the results be applied beyond the mental health counseling context.

Funding This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5C2A02088387).

Data availability The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare no potential conflict of interest concerning the research, authorship, and/or publication of this article.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Afifi, T., & Steuber, K. (2009). The revelation risk model (RRM): Factors that predict the revelation of secrets and the strategies used to reveal them. *Communication Monographs*, 76(2), 144–176. <https://doi.org/10.1080/03637750902828412>
- Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 1(2), 99–112. <https://doi.org/10.1111/j.1468-2958.1975.tb00258.x>
- Brandtzaeg, P. B., & Følstad, A. (2017, November). Why people use chatbots. In: *Proceedings of the 4th International Conference on Internet Science* (pp. 377–392). Springer. https://doi.org/10.1007/978-3-319-70284-1_30
- Burch, N. (2004). Closeness and intimacy. *British Journal of Psychotherapy*, 20(3), 361–372. <https://doi.org/10.1111/j.1752-0118.2004.tb00149.x>
- Burleson, B. R., & Mortenson, S. R. (2003). Explaining cultural differences in evaluations of emotional support behaviors: Exploring the mediating influences of value systems and interaction goals. *Communication Research*, 30(2), 113–146. <https://doi.org/10.1177/0093650202250873>
- Carpenter, A., & Greene, K. (2016). Social penetration theory. *The International Encyclopedia of Interpersonal Communication*, 1–5. <https://doi.org/10.1002/9781118540190.wbeic160>
- Centers for Disease Control and Prevention. (n.d.). Taking Care of Your Emotional Health. Retrieved April 1, 2022, from <https://emergency.cdc.gov/coping/selfcare.asp>
- Chin, Y. C., Su, W. Z., Chen, S. C., Hou, J., & Huang, Y. C. (2018). Exploring users' self-disclosure intention on social networking applying novel soft computing theories. *Sustainability*, 10(11), 1–16. <https://doi.org/10.3390/su10113928>
- Choi, T. R., & Drumwright, M. E. (2021). Choi, T. R., & Drumwright, M. E. (2021). "OK, Google, why do I use you?" Motivations, post-consumption evaluations, and perceptions of voice AI assistants. *Telematics and Informatics*, 62, 101628. <https://doi.org/10.1016/j.tele.2021.101628>
- Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92, 539–548.
- Clark, A. J. (2010). Empathy and sympathy: Therapeutic distinctions in counseling. *Journal of Mental Health Counseling*, 32(2), 95–101. <https://doi.org/10.17744/mehc.32.2.228n116thw397504>
- Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. In J. Norcross (Ed.), *Psychotherapy relationships that work* (2nd ed., pp. 132–152). Oxford University Press.
- Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, 132, 138–161. <https://doi.org/10.1016/j.ijhcs.2019.07.009>
- Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, 4(2), e7785. <https://doi.org/10.2196/mental.7785>
- Fehr, B. (2004). Intimacy expectations in same-sex friendships: A prototype interaction-pattern model. *Journal of Personality and Social Psychology*, 86(2), 265. <https://doi.org/10.1037/0022-3514.86.2.265>
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.
- Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68(4), 712–733. <https://doi.org/10.1093/joc/jqy026>
- Huaman-Ramirez, R., Lunardo, R., & Vasquez-Parraga, A. (2022). How brand self-disclosure helps brands create intimacy with customers: The role of information valence and anthropomorphism. *Psychology & Marketing*, 39(2), 460–477.
- Johnson, R., Hobfoll, S. E., & Zalcberg-Linetzky, A. (1993). Social support knowledge and behavior and relational intimacy: A dyadic study. *Journal of Family Psychology*, 6(3), 266–277. <https://doi.org/10.1037/0893-3200.6.3.266>
- Lee, S., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies*, 103, 1–35. <https://doi.org/10.1016/j.ijhcs.2017.02.005>
- Lee, Y. C., Yamashita, N., Huang, Y., & Fu, W. (2020, April). "I Hear You, I Feel You": Encouraging Deep Self-disclosure through a Chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–12).
- Lepore, S. J., Fernandez-Berrocal, P., Ragan, J., & Ramos, N. (2004). It's not that bad: Social challenges to emotional disclosure enhance adjustment to stress. *Anxiety, Stress & Coping*, 17(4), 341–361. <https://doi.org/10.1080/10615800412331318625>
- Liu, B., & Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychology, Behavior, and Social Networking*, 21(10), 625–636. <https://doi.org/10.1089/cyber.2018.0110>
- Lin, P. Y., Liang, T. P., Huang, H. C., & Li, Y. W. (2021). Design quality, relationship intimacy and continuance intention of mobile apps: An extension to the is success model. *Journal of Electronic Commerce Research*, 22(4), 266–284.
- Levenson, E. A. (1974). Changing concepts of intimacy in psychoanalytic practice. *Contemporary Psychoanalysis*, 10(3), 359–369. <https://doi.org/10.1080/00107530.1974.10745341>
- Mack, H. (2018). Woebot Labs debuts fully AI mental health chatbot via Facebook Messenger. *Mobi Health News*. <https://www.mobihealthnews.com/content/woebot-labs-debuts-fully-ai-mental-health-chatbot-facebook-messenger>

- Malloch, Y. Z., & Zhang, J. (2019). Seeing others receive support online: Effects of self-disclosure and similarity on perceived similarity and health behavior intention. *Journal of Health Communication, 24*(3), 217–225. <https://doi.org/10.1080/10810730.2019.1595226>
- Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research, 26*(4), 323–339.
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Miner, A. S., Milstein, A., & Hancock, J. T. (2017). Talking to machines about personal mental health problems. *JAMA, 318*(13), 1217–1218.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 72–78).
- Portela, M., & Granell-Canut, C. (2017, September). A new friend in our smartphone? Observing interactions with chatbots in the search of emotional engagement. In *Proceedings of the XVIII International Conference on Human Computer Interaction* (pp. 1–7). Psychology, Computer Science. <https://doi.org/10.1145/3123818.3123826>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.
- Reis, H. T., & Franks, P. (1994). The role of intimacy and social support in health outcomes: Two processes or one? *Personal Relationships, 1*, 185–197. <https://doi.org/10.1111/j.1475-6811.1994.tb00061.x>
- Reis, H. T., Lemay, E. P., & Finkenauer, C. (2017). Toward understanding: The importance of feeling understood in relationships. *Social and Personality Psychology Compass, 11*(3), e12308. <https://doi.org/10.1111/spc3.12308>
- Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. W. Duck (Ed.), *Handbook of personal relationships: Theory, research and interventions* (pp. 367–389). Wiley.
- Sanford, K. (2006). Communication during marital conflict: When couples alter their appraisal, they change their behavior. *Journal of Family Psychology, 20*, 256–265.
- Skjuve, M., Haugstveit, I. M., Følstad, A., & Brandtzaeg, P. B. (2019). Help! Is my chatbot falling into the uncanny valley? An empirical study of user experience in human-chatbot interaction. *Human Technology, 15*(1), 30–54.
- Sung, J. Y., Guo, L., Grinter, R. E., & Christensen, H. I. (2007, September). “My Roomba is Rambo”: intimate home appliances. In *International Conference on Ubiquitous Computing* (pp. 145–162). Springer, Berlin, Heidelberg.
- Sweeney, C., Potts, C., Ennis, E., Bond, R., Mulvenna, M. D., O’neill, S., ... & Mctear, M. F. (2021). Can Chatbots help support a person’s mental health? Perceptions and views from mental health-care professionals and experts. *ACM Transactions on Computing for Healthcare, 2*(3), 1–15. <https://doi.org/10.1145/3453175>
- Ting-Toomey, S. (1991). Intimacy expressions in three cultures: France, Japan, and the United States. *International Journal of Intercultural Relations, 15*(1), 29–46. [https://doi.org/10.1016/0147-1767\(91\)90072-O](https://doi.org/10.1016/0147-1767(91)90072-O)
- Vollmer, A. L., Read, R., Trippas, D., & Belpaeme, T. (2018). Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Science Robotics, 3*(21), eaat7111.
- Wang, L., Yan, J., Lin, J., & Cui, W. (2017). Let the users tell the truth: Self-disclosure intention and self-disclosure honesty in mobile social networking. *International Journal of Information Management, 37*(1), 1428–1440. <https://doi.org/10.1016/j.ijinfomgt.2016.10.006>
- Yin, X., Wang, H., Xia, Q., & Gu, Q. (2019). How social interaction affects purchase intention in social commerce: A cultural perspective. *Sustainability, 11*(8). <https://doi.org/10.3390/su11082423>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.