

Anthropomorphic Chatbots for Behavior Change

Seyoung Lee, Ph.D.

Sungkyunkwan University

Department of Media and Communication

LABOR SHORTAGE IN THE POST-PANDEMIC ERA

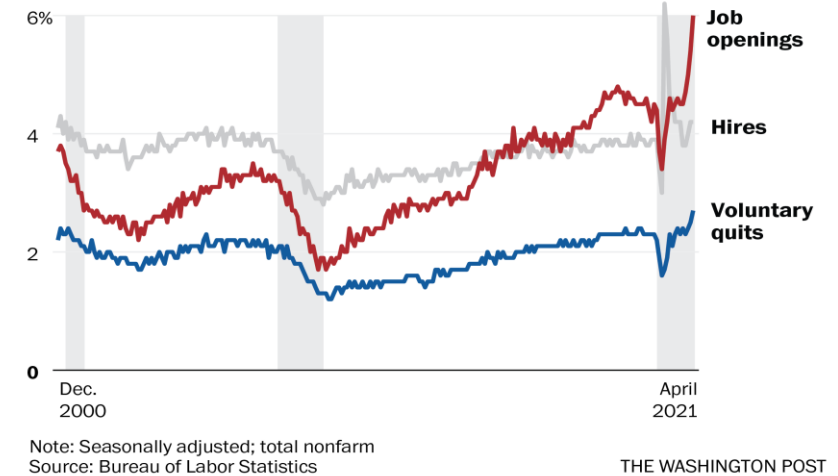
Labor shortage in the post-pandemic era

- Job shock: Demand for labor is historically high
- Competition across industries, ongoing contagion risk
- 40% of adults reported symptoms of anxiety or depression — compared with 11% pre-COVID (Weiner, 2022).
- Burnout and fatigue as the pandemic drags on; “40% of mental health practitioners said they felt burned out last September, and I would imagine that’s only higher now” (van Dam, 2021).
- *“Mental health workers have seen their earnings jump 30% over the course of the pandemic...yet few people have been tempted into the profession”* (van Dam, 2021).

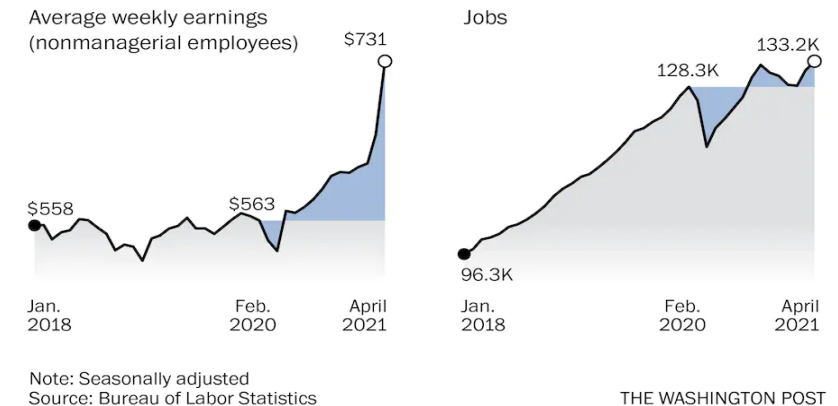
Global job shock: Most affected areas

- *“Nonprofits around the country are reporting significant difficulties retaining staff and filling vacancies”* (National Council of Nonprofits, 2021).
- *Charitable organizations are experiencing a global staffing shortage, which leads to the closing of charitable organizations that are responsible for food, shelter, and mental health of a countless number of beneficiaries including children.*

Job openings and quits hit highest rates on record



Offices of mental health practitioners



LABOR SHORTAGE IN THE POST-PANDEMIC ERA

- [As need increases, nonprofit leaders in Ottawa are exhausted](#), Cassandra Lybrink, *Holland Sentinel*, Sep. 18, 2022.
- [Nonprofits beg for Congress' help with staffing shortages](#), Eleanor Muellor, *Politico*, Sep. 12, 2022.
- ['People will die waiting': America's system for the disabled is nearing collapse](#), Dan Goldberg, *Politico*, Aug. 10, 2022.
- [The Nonprofit Hiring Crisis](#), Jim Rendon, *The Chronicle of Philanthropy*, Jun. 7, 2022.
- [A food pantry's closure means more than lost meals for hundreds of families](#), Kyle Swenson, *Washington Post*, Apr. 18, 2022.
- [Impact of COVID-19 and current restrictions on Ontario's nonprofit sector](#), Ontario Nonprofit Network, Jan. 5, 2022.
- [Nonprofits struggle in the face of COVID-19](#), Nishesh Chalise, Federal Reserve Bank of St. Louis, Dec. 27, 2021.
- [As workers gain pay leverage, nonprofits can't keep up](#), Ben Casselman, *New York Times*, Dec. 23, 2021.
- [Nonprofits struggling to hire staff amid labor shortage](#), Maya Rodriguez, Scripps TV, Nov. 10-11, 2021.

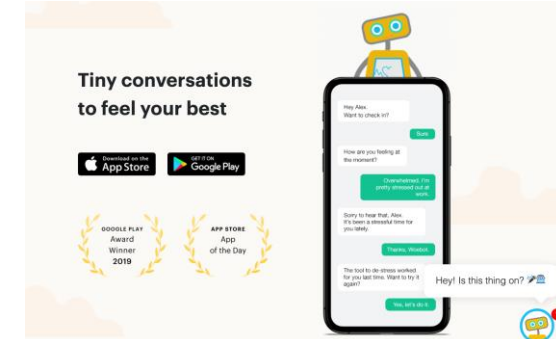
**CHILDREN CAN'T WAIT AND SO DO THE SICK
AND HUNGRY, DESPERATELY IN NEED OF HELP.**



HUMAN-CHATBOT INTERACTION FOR BEHAVIOR CHANGE

The use of human-chatbot interaction to replace social/emotional presence

- Mental health practitioners and charity collectors not only deliver fundraising information but are **actively involved in the persuasion process by forming social and emotional presences** (Chapman et al., 2022; DellaVigna et al., 2012; Harbaugh, 1998; Holländer, 1990; Bénabou & Tirole, 2005).
- Chatbot agents: Youper, Woebot, Moodkits, Moodnotes, Wysa, etc. (Merchant, 2021).
- During the COVID-19 pandemic, 44% high school students reported they persistently felt sad or hopeless (CDC, 2022)
- Privacy, anonymity, ease of access, useful in addressing delicate issues.
- With AI that utilizes machine learning and natural language processing, chatbots serve as companions and offer personalized therapy.
- **Can chatbots replace social/emotional presences of human agents?**
- **When, how, and to what extent?**



Woebot: A supportive mental health chatbot developed by clinical research psychologist Alison Darcy; receives over 2 million messages per week.



Wysa: Utilizes the Facebook Messenger platform as its interface, available in 30 countries, with India, the UK, and the US being its top adopters.

The Computers are social actors (CASA) Paradigm

- People treat computers as social actors and use social rules and norms in an interaction between users and computers (Moon, 2000; Nass & Moon, 2000, Reeves & Nass, 1996).
- Engaging in “mindless” behavior even when such behavior is not necessarily rational in the given context (Moon, 2000)
- The norm of reciprocity would also be valid for artificially generated emotional disclosures.

Humanness cues (Go & Sundar, 2019):

- Identity cues (use of human names and/or identity)
- Conversational cues (mimicking human emotions and languages; empathy, emotional disclosure, etc.)
- Visual cues (use of human figures)

Increased humanness led to:

- Higher sense of companionship (Leite et al., 2013)
- Trust and self-disclosure (Skjuve et al., 2021)
- Positive evaluation, compliance, and several other behavioral outcomes (Go & Sundar, 2019; Paiva et al., 2004; Leite et al., 2013).

Is this always true?

Are there any boundary conditions?

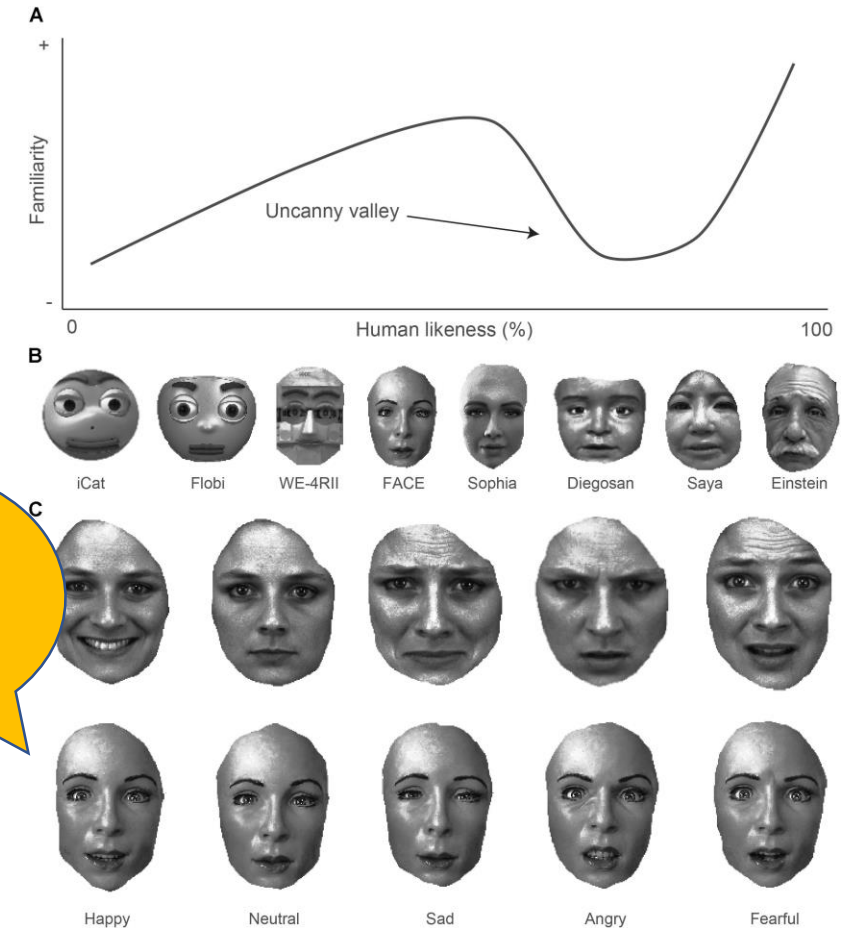
Uncanny Valley Effect

- Humanness perceptions may cause uncomfortable feelings or eeriness and result in negative outcomes (MacDorman, 2019; Mori et al., 2012; Song & Shin, 2022).
- Positive outcomes of humanness drops significantly when it reaches a certain point (Mori et al., 2012).
- Not only AIs with tangible figures but conversational AI agents may experience uncanny valley effect.

When do chatbots' humanness become uncomfortable?

When do we say enough?

Trust me.
I feel what
you feel.



STUDY 1: Effect of ai chatbot emotional disclosure on user satisfaction and reuse intention for mental health counseling: a serial mediation model

Current Psychology, 2022

The Computers are social actors (CASA) paradigm

- People treat computers as social actors and use social rules and norms in an interaction between users and computers (Moon, 2000; Nass & Moon, 2000; Reeves & Nass, 1996).
- People believe that computers' affective state can be influenced by their behavior and emotions.

Emotional exchanges in mental health counseling

- Effective and timely emotional responses are key to increasing user satisfaction and positively impacting mental health and counseling satisfaction (Clark, 2010).
- 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

- Outcomes of emotional disclosure: Increased closeness, trust, enjoyment, and intimacy with a conversational partner (Carpenter & Greene, 2016; Ho et al., 2018).
- Malloch and Zhang (2019) also suggested the effect of emotional disclosure on health behavioral intention.
- Studies suggest that the same effect can be extended to human–computer interaction (Go & Sundar, 2019; Ho et al., 2018; Moon, 2000; Nass & Moon, 2000; Reeves & Nass, 1996).

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 roles of user emotional disclosure intention

The norm of reciprocity: 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).

Uncertainty reduction theory

- Emotional exchanges reduce uncertainty in an early-stage relationship and increase the level of emotional exchange (Berger & Calabrese, 1975).

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 roles of intimacy

- Intimacy has been defined as “the ambiance of closeness” or “shared experience of openness” (Levenson, 1974, pp. 359, 368).
- Emotional disclosure → higher intimacy with a conversational partner (Carpenter & Greene, 2016; Ho et al., 2018) → Lower depression (Reis & Franks, 1994), higher social support satisfaction (Johnson et al., 1993), perceived understanding (Reis et al., 2017), continuous intention (Lin et al., 2021), etc.
- People are motivated by friendly and empathic relationships with computer partners also support the effect of perceived intimacy on outcome variables (Brandtzaeg & Folstad, 2017).

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

- Extant studies have implied the serial mediation of the reciprocity of self-disclosure (Carpenter & Greene, 2016) 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.

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.

Participants

- 348 American adults, recruited from Amazon MTurk in Mar 2022 (158 women and 190 men)

Stimuli

- Online experiment: Two types of chatbots (i.e., factual vs. emotional disclosure) were developed using Dialogflow, an AI-powered conversational platform.
- Factual condition: 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”).
- Emotional disclosure condition: Chatbot counselor disclosed its own emotional experience (e.g., “I felt much better after connecting with others in supportive conversation”).
- Rich response button type option.

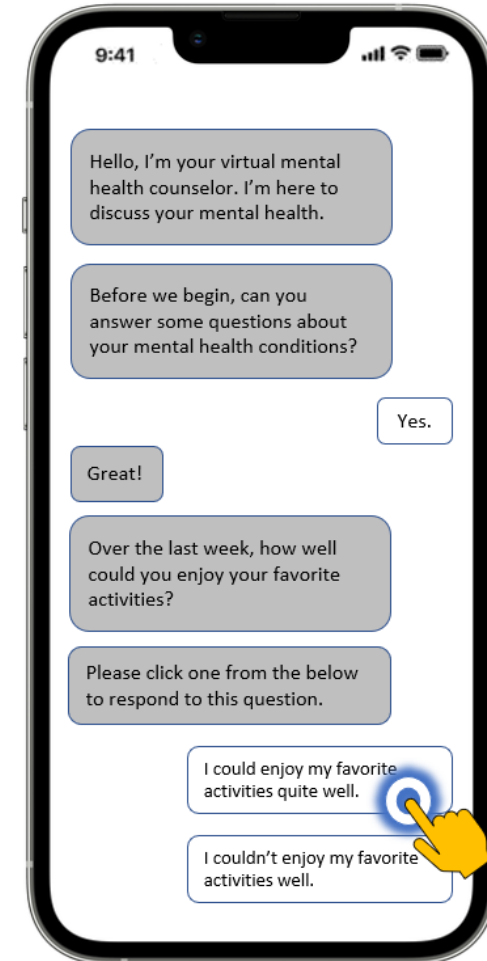


Fig. 1 Button-type chatbot

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

STUDY 1: RESULTS

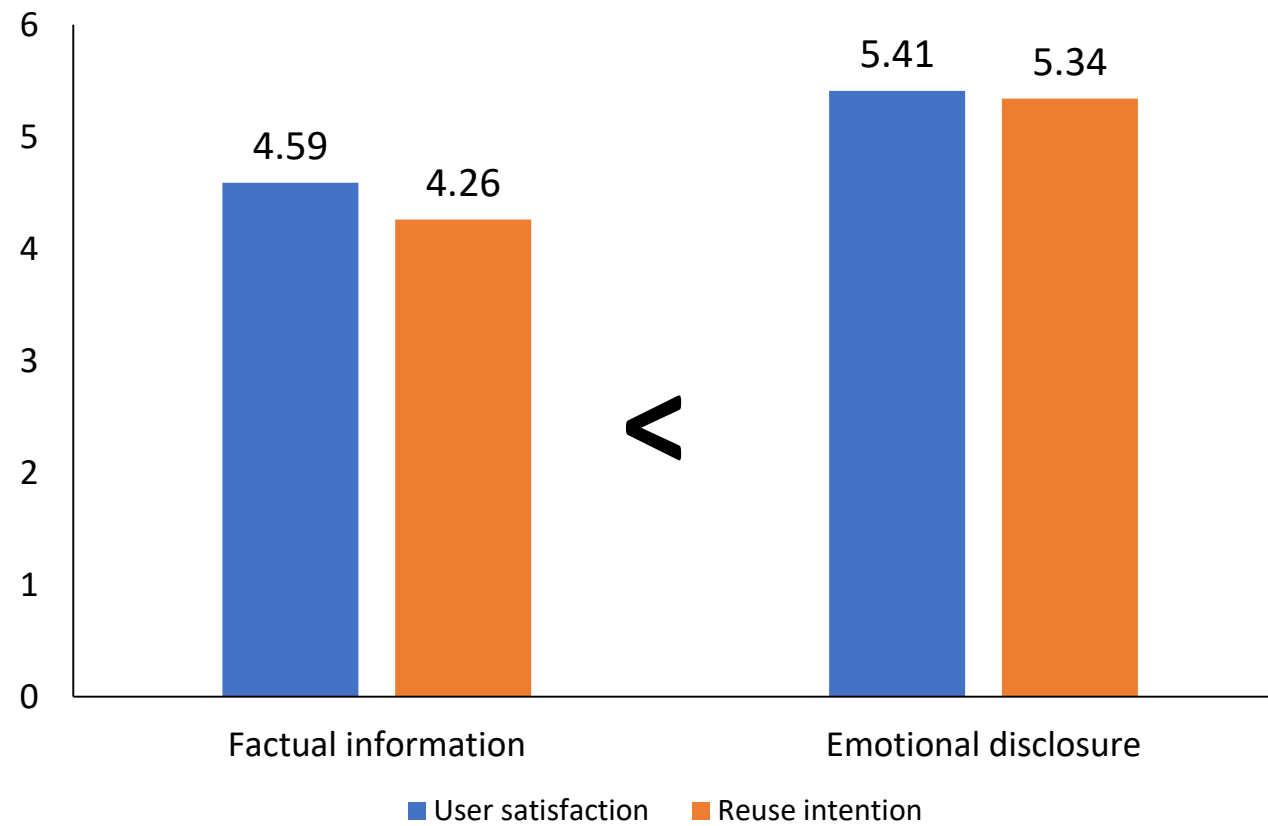


Table 2. Total, direct, and indirect effects of CED on USCC/IRCC (H2, H3, and H4)

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

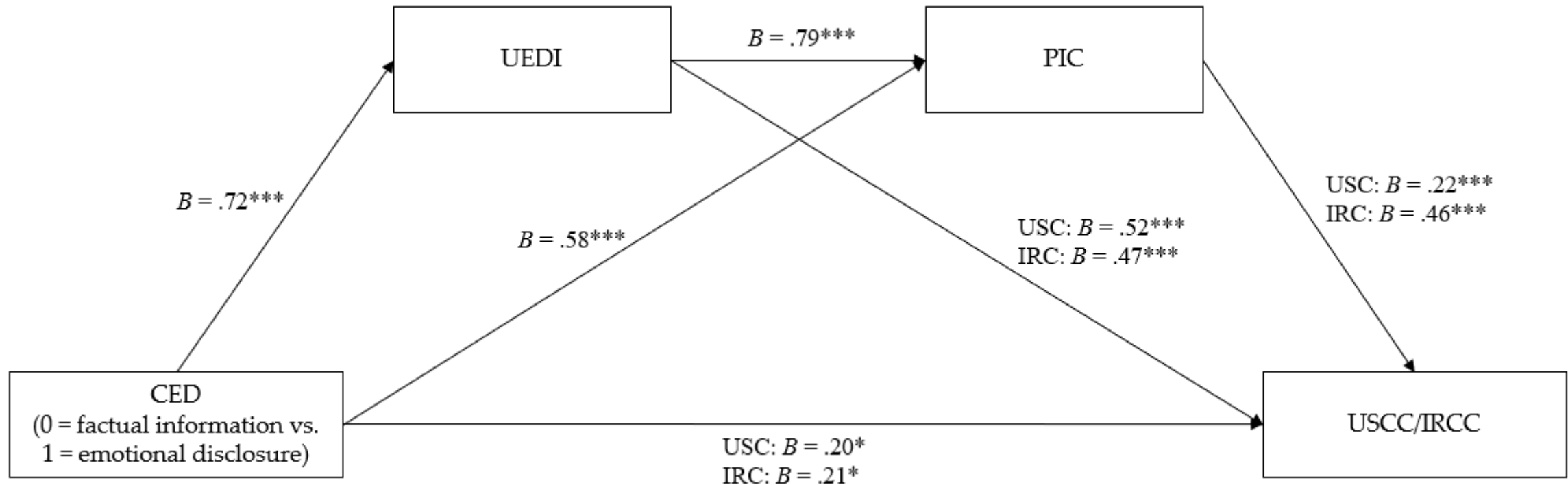


Fig. 2 Unstandardized path coefficients for serial multiple mediation. CED = Chatbot Emotional Disclosure, UEDI = User Emotional Disclosure Intention, PIC = Perceived Intimacy with Chatbot, USCC = User Satisfaction with Chatbot Counseling, IRCC = Intention to Reuse Chatbot Counseling * $p < .05$, * $p < .001$**

- 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.
- The first attempt to test the serial mediation through user emotional disclosure intention and perceived intimacy in the context of chatbot conversation.
- Expands the scope of the theoretical framework on emotional disclosures and exchanges in the context of chatbot conversations and artificial emotions.
- Practical implications for the mental health industry, 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.

STUDY 2: Effect of AI chatbot empathy and identity disclosure on willingness to donate: the mediation of humanness and social presence

Behaviour & Information Technology, 2022

STUDY 2: THEORETICAL PREMISES

The Computers are social actors (CASA) paradigm

- People treat computers as social actors and use social rules and norms in an interaction between users and computers (Moon, 2000; Nass & Moon, 2000, Reeves & Nass, 1996).
- People believe that computers' affective state can be influenced by their behavior and emotions.

Modality-agency-interactivity-navigability (MAIN) model

- Interface cues (e.g., identity, conversational) shape user perceptions and reactions (Go & Sundar, 2019).

Two systems empathy model (Liu and Sundar 2018):

Conversational Cues	Identity Cues
Information only (control)	No name
Cognitive empathy (recognition of feelings)	Robot associated name (i.e., Cybot)
Affective empathy (feel what others feel)	Human-associated name (i.e., Emily)



Higher humanness,
persuasiveness, and
willingness to pay (WTD)
(Go & Sundar, 2019;
Leite et al., 2013)

H1: Chatbot empathy (i.e. cognitive empathy and affective empathy) will yield higher WTD than the information- only condition.

H2: Disclosing a human name (i.e. Emily) will yield higher WTD than a chatbot name condition (i.e. Cybot).

Interaction effect between cues

- Identity cues (i.e. human-associated names) X conversational cues (i.e. mimicking human languages)
- In a humanlike conversation condition, people in a human identity condition showed higher perceived expertise, while the opposite was true in a less humanlike conversation condition (Go & Sundar, 2019).
- A fit between different cues positively affect perception and behavior.

H3: There will be an interaction effect between chatbot empathy and identity disclosure on WTD. The affective empathy conditions with a human name condition will yield higher WTD than the other conditions.

The mediating roles of human likeness and social presence

- Human likeness: The extent to which people feel as if they are interacting with a real human being.
- Social presence: The degree of belief that the other social being coexists and interacts with an individual (Lee & Nass 2005).
- Human likeness & social presence → help assign human-like characteristics → make people become sensitive to their conversational partner's affective states and identity → social facilitation & social pressure → motivation for prosocial behavior, WTD.

H4: Perceived human likeness will mediate the interaction effect of chatbot empathy and identity disclosure on WTD. Specifically, the indirect effect of chatbot empathy on WTD via perceived human likeness will be stronger (vs. weaker) under the human name (vs. chatbot name) condition.

H5: Perceived social presence will mediate the interaction effect of chatbot empathy and identity disclosure on WTD. Specifically, the indirect effect of chatbot empathy on WTD via perceived social presence will be stronger (vs. weaker) under the human name (vs. chatbot name) condition.

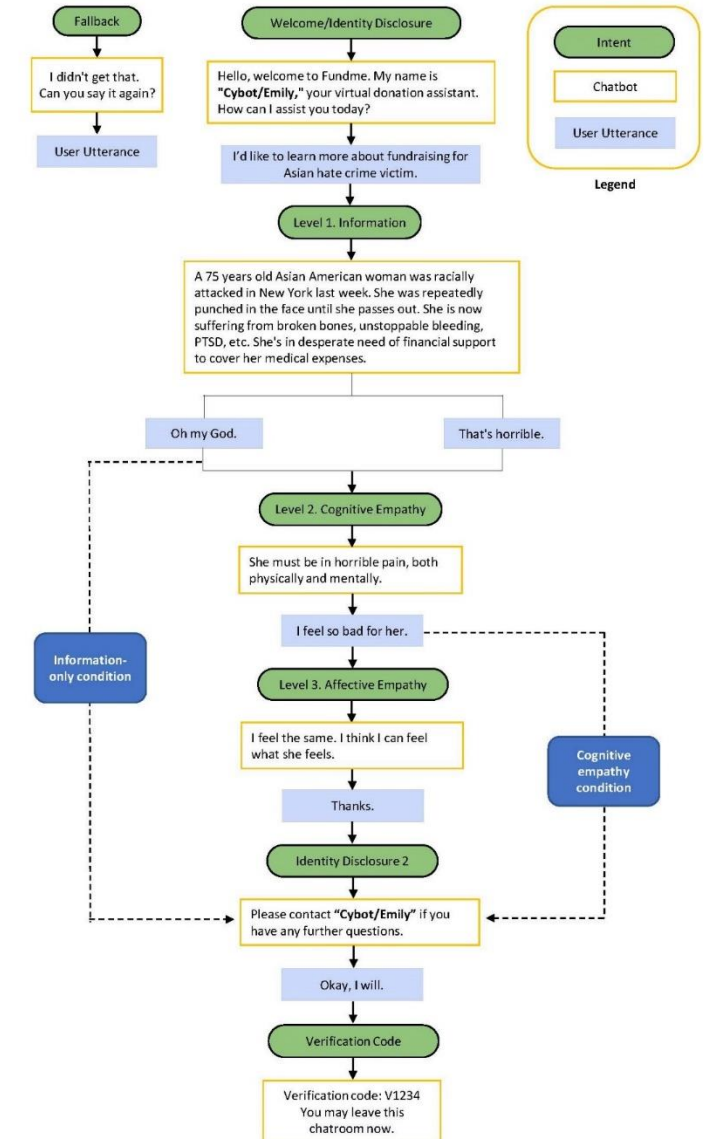
STUDY 2: METHODS

Participants

- 496 American adults, recruited from Amazon MTurk in Oct 2021 (259 women and 237 men)

Stimuli

- Online experiment: A 3 (information-only vs. cognitive empathy vs. affective empathy) × 2 (chatbot vs. human; Cybot vs. Emily) between-subject factorial design.
- Six versions of chatbots were created using Dialogflow, an AI-powered conversational platform.
- Rich response button type option.



Coversational Cues	Identity Cues
Information only (control)	Robot associated name (i.e., Cybot)
Cognitive empathy(recognition of feelings)	
Affective empathy (feel what others feel)	Human-associated name (i.e., Emily)

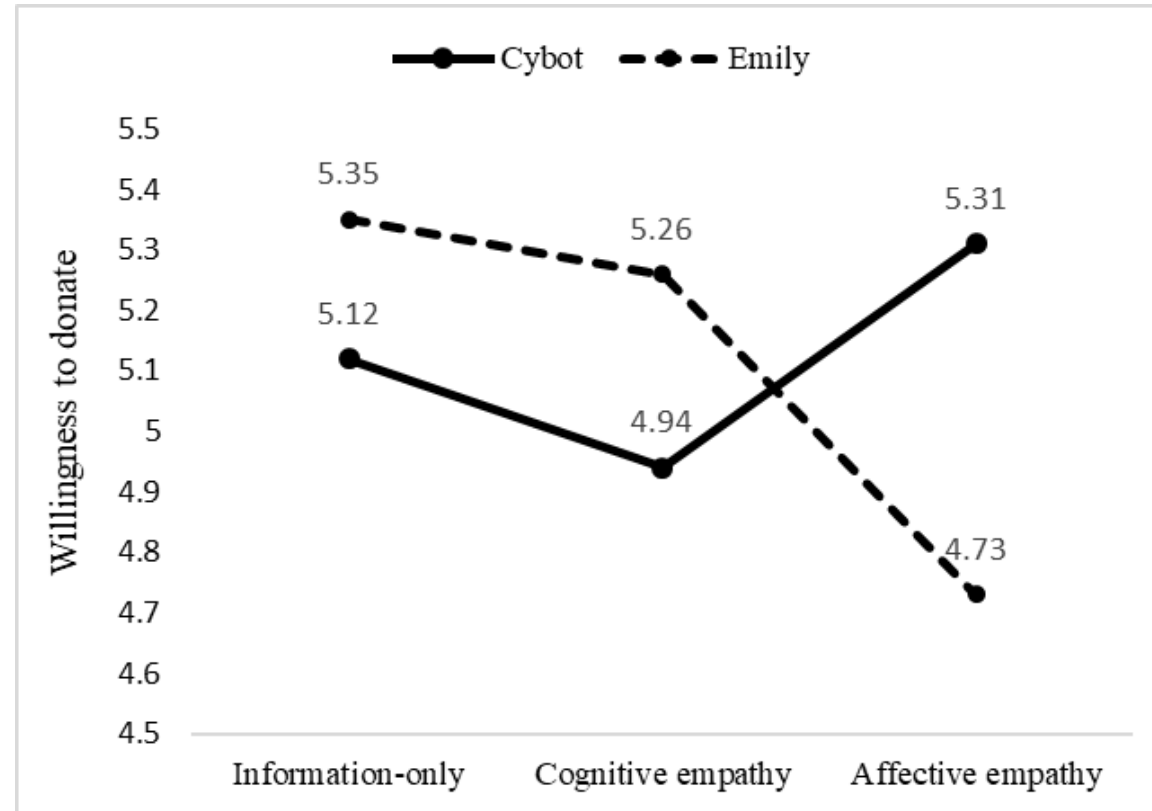
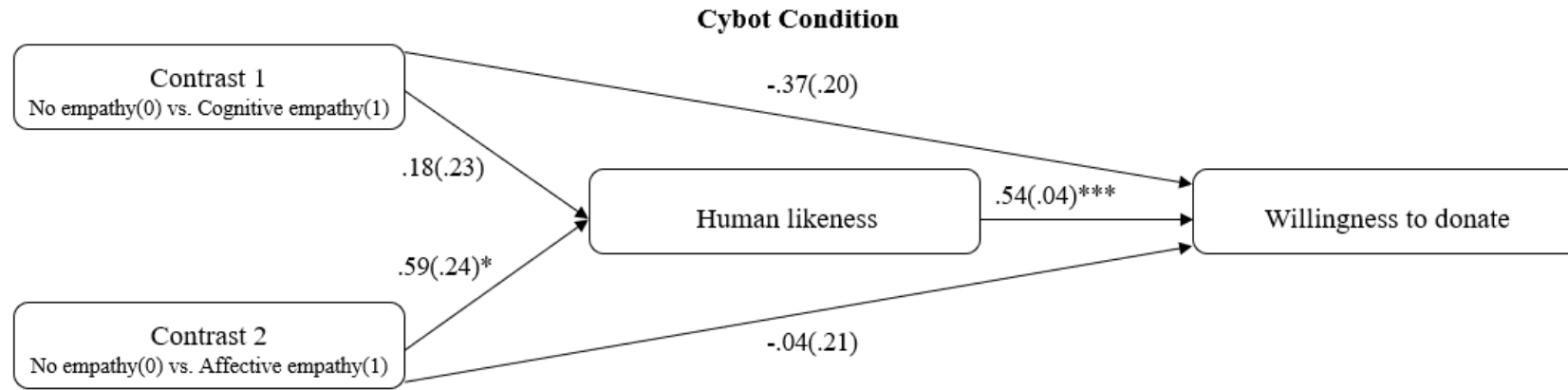


Figure 2. Interaction effect between chatbot empathy and identity disclosure on WTD

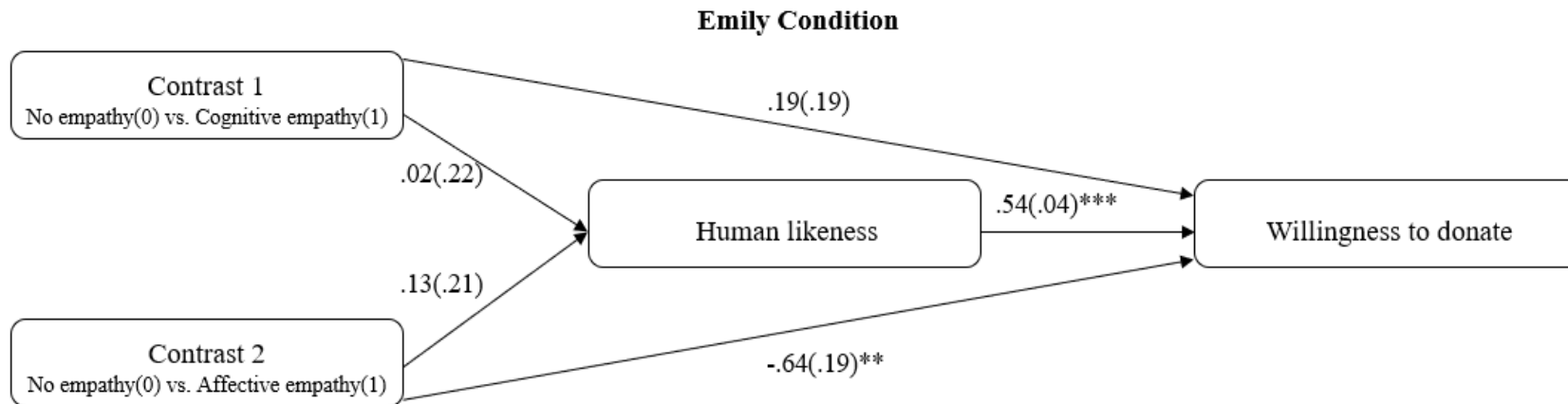
STUDY 2: RESULTS



Indirect effect: Contrast 1 → Human likeness → WTP : $B = .10$, $Boot SE = .12$, $95\% Boot CI = -.13, .32$

Indirect effect: Contrast 2 → Human likeness → WTP : $B = .32$, $Boot SE = .13$, $95\% Boot CI = .06, .58$

Figure 3-1. Unstandardized regression coefficients representing the effect of chatbot empathy on WTD mediated by the human likeness on WTD mediated by human likeness (Cybot condition)



Indirect effect: Contrast 1 → Human likeness → WTP : $B = .01$, $Boot SE = .11$, $95\% Boot CI = -.21, .24$

Indirect effect: Contrast 2 → Human likeness → WTP : $B = .07$, $Boot SE = .12$, $95\% Boot CI = -.16, .30$

Figure 3-2. Unstandardized regression coefficients representing the effect of chatbot empathy on WTD mediated by the human likeness on WTD mediated by human likeness (Emily condition)

STUDY 2: RESULTS

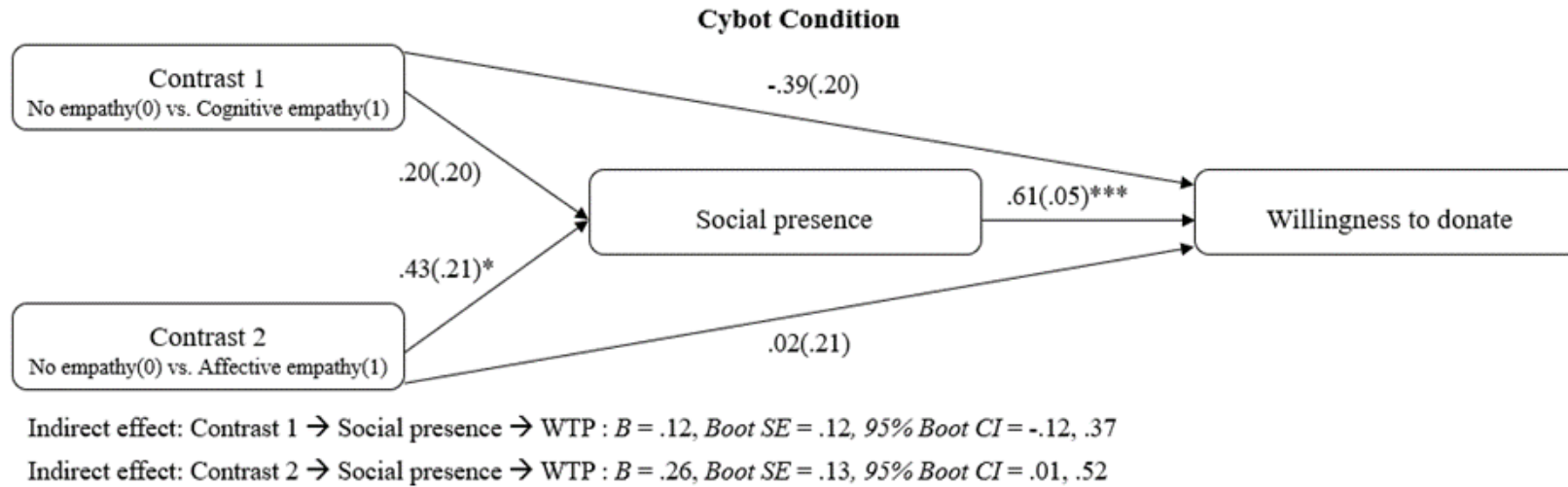


Figure 4-1. Unstandardized regression coefficients representing the effect of chatbot empathy on WTD mediated by the human likeness on WTD mediated by social presence (Cybot condition)

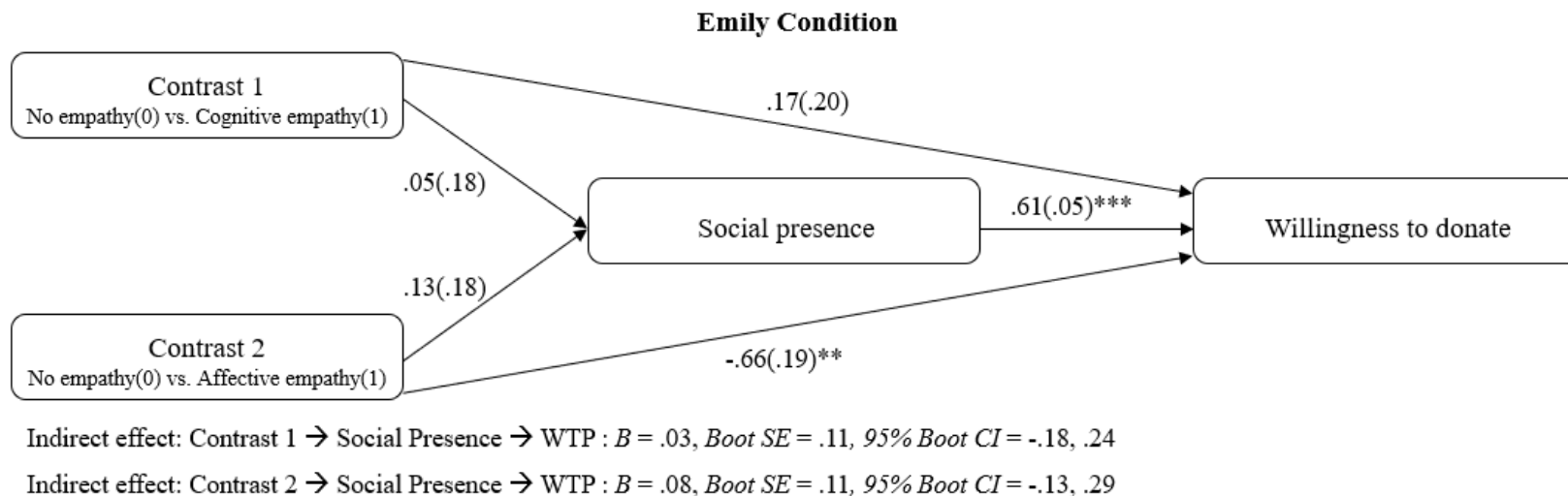


Figure 4-2. Unstandardized regression coefficients representing the effect of chatbot empathy on WTD mediated by the human likeness on WTD mediated by social presence (Emily condition)

- Extending the discussion of chatbot empathy to include chatbot empathy toward a third party (e.g., a victim of a crime) in an emotionally charged situation.
- Broaden the range of possible applications of chatbot empathy from mere advising or information delivery to active persuasion that requires emotional transmission and/or affective interaction.
- Adds to the understanding of the conditions under which the uncanny valley effect occurs in chatbot communications.
- Human identity disclosure disconnects the positive links between chatbot empathy, human likeness/social presence, and WTD.
- Adds to the clarification of the psychological mechanism underlying the relationship between chatbot empathy and user behaviors by testing the mediating roles of human likeness and social presence.

Thank you!