# Human Prejudice to the Healthcare Social Robot and the Impact of Human Personality: An Experiment of the Online Mental Counseling

Runyu SHI
*Warwick Business School*, phd17rs@mail.wbs.ac.uk

Xiang Wan
*University of Warwick*, xinruo2017@gmail.com

# Human Prejudice to the Healthcare Social Robot and the Impact of Human Personality: An Experiment of the Online Mental Counseling

*Completed Research*

**Runyu Shi**
University of Warwick
mandyshiry@gmail.com

**Xiang Wan**
University of Warwick
peter.xiang2020@gmail.com

## Abstract

Robots have vastly influenced the healthcare practices. Ample studies have investigated the use of industrial robots, while leaving social robots an under-explored area. This research investigates human's prejudicial attitude on the social robots that conduct online mental counseling services. 80 participants are recruited, who have been told to receive counseling services either delivered by a human counselor or a robotic counselor. Given that the actual counseling services are all completed by the same human profession with carefully maintained quality consistency, the participants' diverged counseling satisfaction becomes a substantial proxy for their prejudiced attitudes on the robotic counselor. With significant lower counseling satisfaction, participants in the robot group particularly show bias on the "helpfulness", "support", and "inclusion" of counseling services. Furthermore, the personality trait of "openness" is found to significantly narrow the magnitude of human bias on social robot. Overall, this research contributes profound insights to post-Covid healthcare research and practices.

**Keywords**

Social Robot; Online Healthcare; Personality; Mental Counseling

## Introduction

The outbreak of Covid-19 has catalyzed the emergence and prominence of remote healthcare. Enabled by the technological advancements such as artificial intelligence and machine learning, robots have substantially assisted the healthcare practices such as sample analysis, disease tracking, and patient condition monitor and alert. Little attention has been given to the social robot. Furthermore, the Covid-19 pandemic that trembled social activities worldwide triggers an alert for the importance and necessity of online mental healthcare, which can be potentially helped by the social robots. Suggested by Davenport & Kalakota (2019), the social robot is defined as a robotic application with a specific (series of) social-orientated objective and aims to induce and formulate the desired social behaviours. We have seen many instances of social robots such as that designed for treating children with autism (Zhang et al., 2019), assisting elderly care (Abdi et al., 2017), or tranquilizing human mood (e.g. Facebook's Woebot) (Fitzpatrick et al., 2017).

Enabled by the natural language processing technologies, social robot is widely interpreted and recognized as one type of artificial intelligence applications. Therefore, people's attitudes toward such behavior-oriented conversational social tools may widely mirror the conflicting views that made for conventional AI applications. The supporters to social robot may emphasize the capabilities and effectiveness of humanoid technologies in replacing some human-incapable tasks. This stream advocates robots' possibilities in recognizing human language and promptly responding to human needs without showing impatience and bias (Costescu & David, 2014; Demirci, 2018; Skjuve & Brandtzæg, 2018). On the contrary, the opponents to social robot may question the robot's fundamental capabilities in expressing and fulfilling human intimacy. This strand displays untrustworthy on robotic applications to conduct relationship tasks and build appropriate interpersonal connections (Coeckelbergh et al., 2016; Sharkey & Sharkey, 2011). Given such deeply inherited conflicting views, individuals may naturally hold prejudicial opinions to social robot applications without granting any careful evaluation and judgement. Therefore, the fist question asked by this research is: *1) Whether humans display prejudicial attitudes to the service quality of healthcare social robot?*

Besides, informed by the behavior theory, human's sense-making and reaction of other objects and surrounding environment are largely influenced by one's personality (McEachan et al., 2010), this research further proposes the potential effects of human's personality traits in influencing their interpretations and feelings about the social robots. The objective of drawing on personality traits is to uncover some rooted attributes that may distinguish the effectiveness and appropriateness of using social robots to different individuals. This aspect offers particular relevance to the design and deployment of healthcare social robots. Therefore, the second question we ask is: *2) How do individuals' personality influence the magnitude of their prejudice on healthcare social robot?*

## Literature and Hypothesis Development

### *Mental healthcare social robot*

Benefiting from the rapid development of Natural Language Processing (NLP) technology in a recent decade, the AI enabled robot technologies have seen wide applications in speech recognition, text analysis, and intelligent translation. These technological advancements lay the foundation of applying robots in conversational interactions such as conducting robotic mental counseling (Davenport & Kalakota, 2019). Tracing back to the 1970s, a psychologist named Rogers invented a humanoid robot "ELIZA", which is the first-generation robotic counselor designed for the "Personal-Centered Therapy"(Weizenbaum, 1966). The design logic of ELIZA was very straightforward. Patients are required to input their disease' symptoms or describing how they feel. ELIZA will then search user inputs across its pre-installed database and match symptoms to those coded diseases (O'Donnell et al., 2014).

Though the technology no longer sounds like a rocket-science at today, such robot-oriented therapy practices indeed have not received wide implementation yet given the technical difficulty and ethical turbulence. The most socially recognized application of such social robot is FaceFacebook's Woebot. This robot algorithmically read and detect individuals' negative emotions through texts they said, written, and posted. Users can talk to Woebot through text and the Woebot would behave humanoid like a close friend and tranquilize individuals' negative expressions. As a therapeutic tool that directs individuals' feelings and behavior, Woebot performs outstanding services with 87% accuracy in recognizing human mood. Woebot is found to capably identify subtle mood alerts and provide alleviation advice to significantly relieve people's anxiety and reduce the likelihood of mental disorder (Almeida et al., 2017).

The deployment of Facebook's Woebot reveals the possible implementation of the social robot in the mental healthcare sector. The social robot aims to interact with people in the human contextual environment and induce people's desired social emotions and behaviours. As a socially-oriented humanoid robot, the social robot is designed to understand human's expressions of feelings, display sympathy, and convey emotional empathy to humans. With these core features, social robots may present advantages in dealing with those minority social groups or those need special care such as children, the elderly, and the disabled like Alzheimer patients. In this regard, social robots are found to benefit autistic children and reduce their expression of negative feelings such as anxiety, depression, and anger (Diehl et al. 2012). Because the social robots can, to the greatest extent, eliminate "personal" judgement and avoid revealing uncomfortableness or impatience to the treated entities (Wada et al., 2004; Satake et al., 2015).

### *Human prejudice to healthcare social robot*

Allport (1954) and Sampson (1999) once highlight prejudice as a negative attitude based on human's faulty and inflexible generalization. However, their emphasis on negativity, inaccuracy, and incorrect degree of generalization pose high limitations given the complexity of social interactions. In this research, we adopt Jones (1997)'s definition of prejudice as "a positive or negative attitude, judgement or feeling about a person that is generalized from attitudes or beliefs held about the group to which the person belongs". Extending such prejudicial group membership from humans to humanoid social robots, this research proposes that social robot---as an artificial intelligence application that constitutes an instantiation of outgroups to the human begins---will be treated by humans with the prejudicial views.

The outgroup prejudice is emerged from the generalized group categorization that is significantly influenced by people's social tags. Crandall and Stangor (2005) found that people's beliefs and perceptions of others are highly correlated with their own social norms. Individuals may form and accept a prejudicial view if other members within their social context conform to a similar view or broadly agree on its acceptableness (Katz and Braly 1933; Pettigrew 1959). In this regard, we infer that prejudice and perception are often yielded from social constructs and are exemplified within individuals either consciously or unconsciously. First, humans have a mixed feeling of affection-and-fear to the humanoid machines (Ogawa et al., 2011). Humans may display some curiosity about the humanoid objects and hold the deep perception that these robotic objects are algorithm operated entities but not humankind (Robinson et al. 2016). Second, it is not easy to build trust between a robot and a person who has deeply believed his or her needs of intimate human help (Sharkey & Sharkey 2011; Coeckelbergh et al. 2016). This presents huge difficulty of accepting the robotic mental counselor which is conventionally human based. Third, people may face expectancy confusion when interacting with humanoid robots. Especially when robots exhibit anthropomorphic features (e.g., called by a name), people may fall into the "anthropocentric expectancy bias", which implies the human's nature of thinking that "their communication partners will be other humans and may experience expectancy violations when their partners are instead machine interlocutors." (Edwards et al. 2019). Based on such objection, untrustworthy, and unmatched expectancy to social robots that handle interpersonal conversations, we firstly hypothesis that:

*H1: Humans display prejudiced evaluation on healthcare social robot's service quality.*

## *Personality trait as an influential factor*

Informed by the psychology personality theory, people's personality traits can have ubiquitous impacts on people's attitudes, cognitions, and behaviours (Devaraj et al. 2008). Allport (1961) defines personality as "the dynamic organization within the individual of those psychophysical systems that determine his characteristic behavior and thought". It is the personality that forms an individual's essence and thinking patterns. In this study, we adopt Goldberg (1990)'s Five-Factor Model (FFM) to capture individuals' personality differences. The FFM consists of five personality traits: Openness; Conscientiousness; Neuroticism; Extroversion; Agreeableness (Komarraju et al., 2011). As a breakthrough and parsimonious personality model, the FFM possesses high stability to present personality systematically (Costa & McCrae 1992; Briggs 1992; Komarraju et al., 2011). Moreover, it has been approved to sustainably describe human attitudes with retest reliability of a=0.73 (Rammstedt et al. 2013).

The big five factors used by FFM have a specific emphasis on capturing individuals' personality attributes. (a) Openness relates to a person's cognitive style and tolerant to unfamiliar ideas. People with high openness seek abstract thinking and new knowledge, with fewer preferences to traditionalism and conservatism (Kaplan et al., 2015). (b) Conscientiousness refers to individuals' degree of control, organization, and goal-directed motivation and persistence. People of strong conscientiousness tend to apply self-control and delayed satisfaction (Roberts et al.,2009). (c) Extroversion represents individuals' tendency for social interaction and corporation. Extroverts are often enthusiastic about social contact and enjoy people's surroundings (Dumitrache et al., 2018). (d) Agreeableness explains individuals' compassionate and the tendency of interpersonal harmony. This trait also implies people's degree of altruism and value priority between self-interests and external-interests (Graziano & Tobin, 2017). (e) Neuroticism reflects individual's emotional instability, anxiety, and psychological stress, such as unrealistic thoughts, anger, anxiety, and depression (Widiger, 2009).

The personality traits can influence individual's attitudes toward social robot through the following mechanisms. First is the level of social robot's acceptance. This feature is largely determined by the individual's possession of openness. Because this trait implies how individuals may tolerant and cognitively process their views to unfamiliar objects. We hypothesize that:

*H2: The degree of personality openness moderates the magnitude of individual's prejudice on healthcare social robot.*

In addition, given the conversational nature of mental counseling services and general healthcare experiences, individuals' prejudice to the social robot also reflect how would they like to cooperate with other entities and external environment. In this regard, the agreeableness personality trait also plays a role. It implies the degree of the individuals' willingness to perceive the usefulness of social robots rather than the threats of social robots (Huang, 2019; Kaplan et al., 2015). This attributes further indicates how much trust that individuals may give to other social objects (Chopik & Kitayama, 2018; Church, 2016). Building on the agreeableness's social cooperation trend, we further hypothesize that:

*H3: The degree of personality agreeableness moderates the magnitude of individual's prejudice on healthcare social robot.*

# Method

This study is designed as a two-group controlled experiment, where 80 participants are told to be participated in a 30-minutes counseling services conducted either by a human counselor or a robot counselor. We follow the two-group two-time experiment approach (Edwards et al. 2016) that controls individual differences before and after the experiment. The human group is having text-based chat with the professional counselor. The robot group is told to experience the text-based chat with a robotic program that has humanoid nick name and profile. Having the actual same mental counselor behind the scenes to deliver a consistency service quality, the observed differences between the two groups' evaluation on the services satisfaction represent participants' prejudice on social robots, given all other factors equal.

## *Participants*

As our experiment aims to provide healthcare insights to online mental counseling that is particularly relevant to the Covid-19 lockdown circumstance, we recruited 80 Chinese participants studying in the UK universities who have never left the UK since the beginning of lockdown (March 2020). Approximately half of them is female (53.75%). The majority holds a master's degree (86.25%), the remaining 11 participants hold either undergraduate (3.75%) or doctoral degree (10%). We identified in advance that no participant is having mental health disorders nor has taken online mental counseling services before.

As one form of communication, mental counseling service can be understood as a script-based interaction where the counseling consultant leads the conversation based on his or her judgment of the counterparty's different emotional situations (Kellerman 1992; Kollar et al. 2006). Therefore, to ensure the consistent mental counseling service quality, this study invites one professional mental counselor to conduct online counseling activities. The profession is qualified for 2,000 hours of counseling experiences and is certificated by both the Chinese Psychological Association and the American Psychological Association. Involving one single counselor across all observations ensures that the actual service quality is kept consistent in our experiment.

## *Instruments*

We use two instruments to measure the participant's overall satisfaction with online mental counseling. The first scale is the **Session Evaluation Questionnaire (SEQ)**. It consists of a seven-point Likert-type scale consisting of 14 questions, which has been proven to be a reliable tool for measuring people's overall experiences of an object or an activity (Stiles 2002). Besides, given that the mental counseling service also involves high interpersonal intimacy, we adopt the **Inclusion of the Other in the Self (IOS)** to capture the interpersonal closeness. The IOS measures the people's perceived emotional attachment to another individual or object. It is a seven-scale graphic representation proposed by Aron et al. (1992) with 0.85 internal consistency and 0.93 validity. The greater the overlap between "self" and "other", the greater the participant's interpersonal empathy.

Next, to control factors of participant's self-differences before and after the experiment, we use the **Outcome Rating Scale (ORS)** developed by Miller and Duncan(2000). It consists of four items that assess individuals from three aspects: physical and mental health; interpersonal and life social; overall life status. ORS uses the visual analogue scale (10CM) that associates to 0 to 40 score. The more to the right (i.e. higher score), the better the individual's overall life condition is. The average α coefficient of ORS is 0.85, and the retest reliability is about 0.60 (Miller et al., 2003).

Last, we measure the individual's personality using the **Big Five Inventory (BFI)** personality traits. It uses 44 five-point Likert-type questions to measure an individual's personality in five dimensions: openness, conscientiousness; extroversion; agreeableness; and neuroticism. The scale is composed of the typical trait adjectives and situational information. The scale possesses conciseness, clarity, efficiency and time-saving. The scale's internal consistency reliability coefficient is between 0.75 and 0.90 (Fossati et al., 2011; John et al., 1991).

### *Procedures*

Following the two-time measurement approach (Edwards et al. 2016), all participants are told to complete the ORS questionnaire (ORS_pre) that implies their overall life status at the recruitment time. In the meantime, participants also complete the BFI before entering the mental counseling session. Next, over 18 days, 80 participants successionally receive 30-minute mental counseling services. The same counseling profession conducts the counseling sessions in human-human and human-robot groups. For both groups, the counseling profession balances the efforts distributed into listening to participants, rephrasing participants' responses, asking follow-up questions, and providing suggestions (Leite et al., 2013; Shimada et al., 2012; Tanaka et al., 2007). Immediately after each counseling session, the participant will complete the ORS questionnaire (ORS_post) again to capture the possible personal differences from the recruitment date to the counseling date. Besides, participants are asked to evaluate their satisfaction with counseling experience using SEQ and their interpersonal intimacy with the consultant (or robot) using IOS. At the end of the experiment, all participants are thanked with five-pound cash.

## Analysis

### *Group differences*

Table 1 reports the summary statistics of 80 participants on all measurement dimensions. The satisfaction rating and observations' age are adjusted with logarithm transformation. The pairwise correlation between every two measurements is included. According to the results, the individual's overall life status after the experiment (ORS_post) is highly correlated to that before the experiment (ORS_pre). This shows relatively stable conditions of the engaged participants. Besides, the significant correlation between individuals' satisfaction with openness (PO) and agreeableness (PA) shows that a person's personality traits may associate to their evaluation of the mental counseling experiences.

**Table 1. Summary Statistics and Pairwise Correlation**

| Variables | Mean | Std. Dev. | min | Med | max | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Satisfy_log | 4.30 | 0.23 | 3.47 | 4.34 | 4.64 | 1.000 | | | | | | | | |
| (2) Age_log | 3.14 | 0.05 | 3.04 | 3.14 | 3.33 | 0.048 | 1.000 | | | | | | | |
| (3) ORS_pre | 25.42 | 3.40 | 20 | 25 | 32 | -0.036 | 0.018 | 1.000 | | | | | | |
| (4) ORS_post | 27.23 | 4.35 | 17 | 28 | 35 | 0.078 | 0.016 | 0.886* | 1.000 | | | | | |
| (7) PO | 38.21 | 6.42 | 15 | 39 | 49 | 0.771* | 0.116 | 0.028 | 0.092 | 1.000 | | | | |
| (8) PC | 29.08 | 8.19 | 14 | 30 | 44 | 0.155 | 0.044 | 0.119 | 0.171 | 0.131 | 1.000 | | | |
| (5) PE | 25.76 | 8.10 | 12 | 26 | 40 | 0.041 | 0.002 | 0.062 | 0.094 | 0.098 | 0.483 | 1.000 | | |
| (6) PA | 34.08 | 6.21 | 19 | 34 | 45 | 0.667* | 0.153 | -0.058 | 0.022 | 0.131 | 0.075 | 0.157 | 1.000 | |
| (9) PN | 26.95 | 7.00 | 15 | 27.5 | 40 | 0.036 | 0.279 | 0.091 | 0.143 | 0.130 | 0.155 | 0.069 | -0.002 | 1.000 |

Note. The number of observations is 80 for all variables. And the significant level of correlation is starred if it is significant at 0.001significance level. PO, PC, PE,PA, PN are representing the personality's trait of openness; conscientiousness; extroversion; agreeableness; and neuroticism.

Table 2 shows the participants' biased evaluation on social robot counseling quality. According to the results, individuals that assigned to the robot group have significant lower satisfaction score than those in the human group. Given that all other conditions such as ORS ratings, personality traits, and service quality are generally balanced, such rating differences significant at 0.05 level confirms participants' prejudicial evaluation on social robot's mental counseling service quality. Furthermore, Table 3 provides a detailed explanation of the participants' rating rationales. Aspects that particularly receive lower ratings are "helpful", "supportive", "inclusion", and "no scruple". These dimensions greatly mirror human's judgment criteria of human interactions, thus indicates individuals' bias from unmatched expectation when interact with social robots.

**Table 2. Sample difference test between AI group and Human group**

| Variables | Obs | | Mean | | Diff | St_Err | t | p |
|---|---|---|---|---|---|---|---|---|
| | Robot | Human | Robot | Human | | | | |
| (1) log_Satisfaction | 40 | 40 | 4.242 | 4.365 | -0.124 | .051 | -2.45 | 0.017** |
| (2) log_Age | 40 | 40 | 3.143 | 3.135 | .007 | .012 | 0.6 | 0.554 |
| (3) ORS1 | 40 | 40 | 25.45 | 25.4 | 0.050 | 0.766 | 0.05 | 0.948 |
| (4) ORS2 | 40 | 40 | 27.325 | 27.15 | 0.175 | 0.978 | 0.2 | 0.859 |
| (5) Extroversion | 40 | 40 | 25.25 | 26.275 | -1.025 | 1.819 | -0.55 | 0.575 |
| (6) Agreeableness | 40 | 40 | 32.975 | 35.2 | -2.225 | 1.373 | -1.6 | 0.11 |
| (7) Openness | 40 | 40 | 37.925 | 38.5 | -0.575 | 1.443 | -0.4 | 0.692 |
| (8) Conscientiousness | 40 | 40 | 27.85 | 30.325 | -2.475 | 1.821 | -1.35 | 0.178 |
| (9) Neuroticism | 40 | 40 | 25.575 | 28.325 | -2.750 | 1.544 | -1.8 | 0.079* |

Note. *0.01 **0.05 ***0.001

**Table 3. Sample difference test on 15 satisfaction dimensions**

| | Robot | Human | Diff | St_Err | t | p |
|---|---|---|---|---|---|---|
| S1 Easy | 4.775 | 5.9 | -1.125 | .374 | -3 | 0.004** |
| S2 Good | 5.1 | 5.275 | -.175 | .384 | -.45 | 0.649 |
| S3 Valuable | 4.95 | 5.425 | -.475 | .381 | -1.25 | 0.215 |
| S4 Deep | 4.575 | 5.475 | -.9 | .394 | -2.3 | 0.025** |
| S5 Relaxed | 4.925 | 5.525 | -.6 | .401 | -1.5 | 0.139 |
| S6 Full | 4.75 | 5.575 | -.825 | .368 | -2.25 | 0.028** |
| S7 Pleasant | 4.6 | 5.4 | -.8 | .402 | -2 | 0.05* |
| S8 Powerful | 4.75 | 5.2 | -.45 | .426 | -1.05 | 0.294 |
| S9 Special | 5.05 | 5.4 | -.35 | .394 | -.9 | 0.377 |
| S10 Smooth | 4.825 | 5.775 | -.95 | .367 | -2.6 | 0.011** |
| S11 Comfortable | 4.8 | 4.95 | -.15 | .393 | -.4 | 0.704 |
| S12 Helpful | 4.375 | 5.6 | -1.225 | .276 | -4.45 | 0.000*** |
| S13 Supportive | 4.175 | 5.4 | -1.225 | .316 | -3.9 | 0.000*** |
| S14 No Scruple | 5.925 | 3.375 | 2.55 | .278 | 9.2 | 0.000*** |
| S15 Inclusion | 4.35 | 5.475 | -1.125 | .283 | -3.95 | 0.000*** |

Note. The number of observations is 80 for all fifteen dimensions. *0.01 **0.05 ***0.001

## *The moderating effects of personality traits*

To quantitatively investigate how participant's personality traits may influence their degree of prejudice on social robot, we further apply Equation 1 with our samples. The $Y_i$ is the dependent variable, which is the satisfaction score of the counseling services rated by observation $i$. The $G_i$ is the group identifier that equals to one if the observation $i$ is from the robot group and equals to zero if it is in human group. The γ represents a group of five coefficients that corresponding to five personality traits of observation $i$. The interaction term $G_i * \sum_{i=1}^{5} X_i$ indicates, whether robot group's prejudice is reinforced or cannibalized by participant's personality traits. θ is the control variables such as observation's pre-experiment life status, gender, age and education. And the ε captures error.

$$Y_i = \alpha + \beta G_i + \gamma \sum_{i=1}^{5} X_i + \delta\, G_i * \sum_{i=1}^{5} X_i + \theta + \varepsilon \qquad \text{Equation 1}$$

Results are reported in Table 4. Column (1) is the baseline model where satisfaction is predicted based on observations' assigned groups. In column (2), the Big Five Inventory traits are added. Results reveal that observations in the robot group generally show a 9% lower satisfaction score. Keeping all factors equal, the one unit increase in the observation's agreeableness and openness level will increase the satisfaction score by 1.4% and 2.2%, respectively. In column (3), we add control variables, and the implications remain valid. Column (4) includes the interaction of the robot group and individuals' personality traits. Results reveal that one unit increase in their personality' openness can reduce their degree of prejudice by around 1.7% for participants in the robot group.

**Table 4. Regression modelling of participants' rating score on mental counseling satisfaction**

| Variable | (1) Satisfication_log | (2) Satisfication_log | (3) Satisfication_log | (4) Satisfication_log |
|---|---|---|---|---|
| Robot_group | -0.124(-2.44) * | -0.090 (-3.20) ** | -0.073(-2.47) * | -0.586(-3.01) ** |
| Personality_O | | 0.022(9.01) *** | 0.023(9.27) *** | 0.013(3.95) *** |
| Personality_C | | 0.000(0.13) | 0.001(0.54) | 0.001(0.63) |
| Personality_E | | -0.001(-0.29) | -0.001(-0.96) | 0.001(0.32) |
| Personality_A | | 0.014(6.62) *** | 0.014(6.73) *** | 0.013(5.29) *** |
| Personality_N | | -0.003(-1.70) | -0.003(-1.45) | -0.002(-0.67) |
| Robot_group * PO | | | | 0.017(3.61) *** |
| Robot _group * PC | | | | -0.001(-0.30) |
| Robot _group * PE | | | | -0.003(-0.81) |
| Robot _group * PA | | | | -0.002(-0.47) |
| Robot _group * PN | | | | 0.001(0.37) |
| ORS_pre | | | -0.003(-0.81) | -0.003(-0.72) |
| Age_log | | | -0.684(-1.58) | -0.840(-2.18) * |
| Gender_male | | | -0.035(-1.17) | -0.032(-1.12) |
| Education_pg | | | -0.011(-0.18) | 0.029(0.44) |
| Education_phd | | | 0.119(1.18) | 0.135(1.32) |
| R2 | 0.071 | 0.747 | 0.765 | 0.808 |
| F statistics | 5.95** | 40.63*** | 26.95*** | 37.13*** |
| N | 80 | 80 | 80 | 80 |

Note. Robust standard errors are included in parenthesis. *o.o1 **o.o5 ***o.oo1

# Discussion

## *Prejudice of healthcare social robot*

According to results, for both human and robot group, the overall mental conditions that observations experience are significantly improved after the experiment (ORS_post) compared to that of before the experiment. It shows the effectiveness and helpfulness of mental counseling services in general. However, regarding the participant's evaluation of their counseling experiences. The robot group shows a significantly lower score than the human group. Given that the healthcare practices and service quality are levelled between two groups, such satisfaction divergence can be interpreted as to the participant's prejudice to the social robot's service quality.

Going further to the detailed criteria that are used in measuring SEQ and IOS, the participants from the robot group and human group are found to display no significant differences in term of "good", "relaxed", "powerful", and "comfortable" (S2, S5, S8, S11). These dimensions imply several characteristics where the robot practices can be equally matched to human conversation expectations. However, the robot group has a particularly lower score on "easiness", "deep", "full", and "smooth" (S1, S4, S6, S10). It implies that, compared with the human group, the robot group generally express negative views that the conversation is more difficult, shallow, and less smooth. Given the experimental-fixed counseling quality, such differences indicate that users may rate the quality of conversation with the stereotyped impression of robot applications. Participants tend to follow the view that AI robots are just programmed objects that cannot provide empathetic responses in any means. This further leads them to presumably think that the interactions with the social robot can be difficult and lack of depth. Besides, the robot group also shows a significantly lower score in terms of "helpfulness", "support", "inclusion" (S12, S13, S15). Even if participants in the robot group are receiving the same level of counseling services, they tend to think that the robots are not so supportive and intimate. This finding reflects the unmatched expectancy when conventional human-centered services are replaced by robotic programs.

### *Personality trait as moderating factors*

First, participants in the robot and human group do not display significant and systematic differences in their personality trait, which to some extent simulate a randomized sample. Second, the analysis approves the impacts of openness and agreeableness on participants' overall satisfaction with counseling experiences. Individuals with high openness prefer abstract thinking and new venture. Therefore, they are generally more open to a new type of social practices. Individuals with high agreeableness tend to be considerate, willing to think from the other side's perspective, and often have empathetic behaviours.

Regarding the personality's moderating effects on individual prejudice, this research found that "openness" is an effective moderator in reducing the magnitude of bias. And the "agreeableness" does not play a significant role in this regard. Because the social robot is, in general, a novel AI application that replaces the traditional human counselors, it is more easily for individuals with higher openness level to accept such innovation. In other words, given the robot group's general prejudice, the degree of such biased attitudes can be reduced if participants possess higher openness personality trait.

## Contribution and limitation

As an interdisciplinary study that combines information systems with psychology, this research contributes to the existing literature in several ways. First, this research uncovers the prejudice that held by people on healthcare social robots and several measurement aspects. It brings a new understanding of designing social robots in healthcare practices. Second, this research also connects robot deployment with individuals' personality. It proposes the users' personality trait as an influential factor that influences their evaluation of social robot interactions. Particularly, people with higher openness are more easily to be educated to accept social robot in healthcare. These findings can help healthcare practitioners to understand better and target users when design and deploy social robots. The social robot should be particularly promoted of the supportive and inclusive features in its design. Besides, social robots are more likely to win early market if it can possibly target people with high openness personality trait.

This research complements past studies that have examined the relationship between people's personality and robot-human interaction in two ways. First, unlike Devaraj et al. (2008), which focuses on people's acceptance level of general technology, this research is established on an alternative interaction medium between humans and robots. In this sense, robots are not subordinated technologies operated by humans, but an essential force that forms human social life. Second, Muller and Richert (2018) once shed lights on the immense human-robot interaction. However, they do not prevent human's subjective perceptions in their methodological design. Participants involved are mostly college students who majored in technical subjects. Such research sample tend to have frequent contact and familiarity with robotic technologies. Besides, participants are well informed of the experiment purposes by given the Negative Attitudes toward Robots Scale (NARS) questionnaire, which may induce participants to consciously be aware of their negative feelings to the robot.

For future research, the deeper knowledge of participants experiences can be investigated through text analysis, field observation or interview. Besides, participants recruited in this experiment are having the same country background. Further research may engage a diverse cultural background and cross validate the prejudice by take account of different cultural background. Besides, this study considers gender, age, and educational background as control variables. It would be very interesting to explore more biological or social factors that may influence prejudice on the social robot.

## REFERENCES

Abdi, J., AI-Hindawi, A., Ng,T., Vizcaychipi, MP. 2017. “Scoping Review on the use of Socially Assistive Robot Technology in Elderly Care,” BMJ Open 8:e018815. doi:10.1136/ bmjopen-2017-018815.

Allport, G.W. 1954. The nature of prejudice. Garden City, NY: Doubleday.

Allport, G.W. 1961. Pattern and Growth in Personality. Holt, Reinhart & Winston.

Almeida, H., Briand, A., & Meurs, M.J. 2017. “Detecting Early Risk of Depression from Social Media User-generated Content,” *Paper presented at the Conference and Labs of the Evaluation Forum (Working Notes).*

Aron, A., Aron, E. N., & Smollan, D. 1992. "Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness," *Journal of Personality and Social Psychology (63*:4), pp.596-612.

Briggs, S.R. 1992. "Assessing the Five-Factor Model of Personality Description," *Journal of Personality* (60:2), pp.253-293.

Chopik, W. J., & Kitayama, S. 2018. "Personality Change Across the Life Span: Insights from a Cross - Cultural, Longitudinal Study," *Journal of personality (86*:3), pp.508-521.

Church, A. T. 2016. "Personality Traits Across Cultures," *Current Opinion in Psychology (8*), pp. 22-30.

Coeckelbergh, M., Pop, C., Simut, R., Peca, A., Pintea, S., David, D., & Vanderborght, B. 2016. "A Survey of Expectations about the Role of Robots in Robot-assisted Therapy for Children with ASD: Ethical Acceptability, Trust, Sociability, Appearance, and Attachment," *Science and Engineering Ethics (22*:1), pp.47-65.

Costa Jr, P. T., & McCrae, R. R. 1992. "Four Ways Five Factors are Basic," *Personality and Individual Differences (13*:6), pp. 653-665.

Costescu, C. A., & David, D. O. 2014. "Attitudes toward Using Social Robots in Psychotherapy," *Transylvanian Journal of Psychology (15*:1).

Crandall,C.S. & Stangor,C. (Eds.) 2005. "Conformity and Prejudice" Malden, MA: Blackwell.

Davenport, T., & Kalakota, R. 2019. "The Potential for Artificial Intelligence in Healthcare," *Future healthcare journal (6*:2), pp.94-98..

Demirci, H. 2018. "User Experience Over Time with Conversational Agents Case Study of Woebot on Supporting Subjective Well-Being," Turkey: Middle East Technical University.

Devaraj,S., Easley,RF., Crant, M. 2008. "How Does Personality Matter? Relating the Five-Factor Model to Technology Acceptance and Use," *Information Systems Research* (19:1), pp.93-105.

Diehl, J. J., Schmitt, L. M., Villano, M., & Crowell, C. R. 2012. "The Clinical Use of Robots for Individuals with Autism Spectrum Dsorders: A Critical Review," *Research in Autism Spectrum Disorders (6*:1), pp. 249-262.

Dumitrache, C. G., Rubio, L., & Rubio-Herrera, R. 2018. "Extroversion, Social Support and Life Satisfaction in Old Age: A Mediation Model," *Aging & Mental Health, (22*:8), pp. 1069-1077.

Edwards, A., Edwards, C., Westerman, D., Spence, PR. 2019. "Initial Expectations, Interactions, and Beyond with Social Robots," *Computer in Human Behavior* (90), pp.308-314.

Edwards, C., Edwards, A., Spence, P., & Westerman, D. 2016. "Initial Interaction Ex-pectations with Robots: Testing the Human-to-Human Interaction Script," *Communication Studies* (67), pp. 227–238.

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), e19.

Fossati, A., Borroni, S., Marchione, D., & Maffei, C. 2011. "The Big Five Inventory (BFI): Reliability and Validity of its Italian Translation in Three Independent Nonclinical Samples," *European Journal of Psychological Assessment (27*:1), pp. 50–58.

Goldberg, L. R. 1990. "An alternative 'Description of Personality': The Big-Five Factor Fracture," *Journal of personality and social psychology* (59:6), pp.1216-1229.

Graziano, W. G., & Tobin, R. M. 2017. "Agreeableness and the Five Factor Model," *The Oxford handbook of the five factor model*, pp. 105-132.

Huang, C. 2019. "Social Network Site Use and Big Five Personality Traits: A meta-Analysis," *Computers in Human Behavior(97)*, pp. 280-290.

John, O. P., Donahue, E. M., & Kentle, R. L. 1991. "The Big-Five Inventory-Version 4a and 54" Berkeley, CA: Berkeley Institute of Personality and Social Research, University of California.

Jones JM.1997. "Prejudice and Racism". 2nd ed McGraw-Hill; New York.

Kaplan, S. C., Levinson, C. A., Rodebaugh, T. L., Menatti, A., & Weeks, J. W. 2015. " Social Anxiety and the Big Five Personality Traits: The Interactive Relationship of Trust and Openness," *Cognitive Behaviour Therapy (44*:3), pp.212-222.

Katz, D., & Braly, K.". 1933. "Racial Stereotypes of One Hundred College Students," *Journal of Abnormal and Social Psychology* (28), pp. 280-290.

Kellerman, K. L. 1992. "Communication: Inherently Strategic and Primarily Automatic," *Communication Monographs* (59), pp. 288–300.

Kollar, I., Fischer, F., & Hesse, F. W. 2006. "Collaboration Scripts - A Conceptual analysis," *Educational Psychology Review* (18), pp.159–185.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. 2011. "The Big Five Personality Traits, Learning Styles, and Academic Achievement," *Personality and individual differences, (51*:4), pp. 472-477.

Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., & Paiva, A. 2013. "The Influence of Empathy in Human–Robot Relations," *International Journal of Human-Computer Studies (71*:3), pp.250-260.

McEachan, RC., Sutton,S., Myers,L. 2010. "Mediation of Personality Influences on Physical Activity within the Theory of Planned Behaviour," *Journal of Health Psychology* (15:8), pp 1170-1180.

Miller, S. D., & Duncan, B. 2000. The outcome rating scale. Chicago, IL: authors.

Miller, S. D., Duncan, B., Brown, J., Sparks, J., & Claud, D.2003. "The Outcome Rating Scale: A Preliminary Study of the Reliability, Validity, and Feasibility of a Brief Visual Analog Measure," *Journal of brief Therapy (2*:2), pp. 91-100.

Muller, SL., Richer,A. 2018. "The Big-Five Personality Dimensions and Attitudes Towards Robots: A Cross Sectional Study," *The 11th Pervasive Technologies Related to Assistive Environments Conferences, Corfu, Greece*.

O'Donnell, K., Dorsey, S., Gong, W., Ostermann, J., Whetten, R., Cohen, J. A., Whetten, K. 2014. "Treating Maladaptive Grief and Posttraumatic Stress Symptoms in Orphaned Children in Tanzania: Group - based Trauma - Focused Cognitive–behavioral Therapy," *Journal of traumatic stress (27*:6), pp.664-671.

Ogawa, K., Nishio, S., Koda, K., Balistreri, G., Watanabe, T., & Ishiguro, H. 2011. "Exploring the Natural Reaction of Young and Aged Person with Telenoid in a Real World," *Journal of Advanced Computational Intelligence and Intelligent Informatics* (15:5), pp. 592-597.

Pettigrew, T.F. 1959. "Regional Differences in Anti-Negro Prejudice," *Journal of Abnormal and Social Psychology*(59), pp.28-36.

Rammstedt, B., Kemper, C.J. Klein,M.C. Beierlein, C Kovaleva, A., 2013. "10 Item Big Five Inventory (BFI-10),"*Methoden, Daten, Analyse (*7:2), pp. 233–249.

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. 2009. Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (p. 369–381). The Guilford Press.

Robinson, H., Broadbent, E., & MacDonald, B. 2016. "Group Sessions with Paro in a Nursing Home: Structure, Observations and Interviews," *Australasian journal on ageing, (35*:2), pp. 106-112.

Sampson, E.E. 1999. "Dealing with Differences: An Introduction to the Social Psychology of Prejudice," Fort Worth: Harcourt Brace.

Satake, S., Hayashi, K., Nakatani, K., & Kanda, T. 2015. "Field Trial of an Information-Providing Robot in a Shopping Mall," *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Sharkey, A., & Sharkey, N. 2011. "Children, the Elderly, and Interactive Robots," *IEEE Robotics & Automation Magazine (18*:1), pp. 32-38.

Shimada, M., Kanda, T., & Koizumi, S. 2012. "How Can a Social Robot Facilitate Children's Collaboration? *International Conference on Social Robotics*.

Skjuve, M., & Brandtzæg, P. B. 2018. "Chatbots as a New User Interface for Providing Health Information to Young People," *Youth and news in a digital media environment–Nordic-Baltic perspectives*.

Stiles, W. B. 2002. "Session evaluation questionnaire: Structure and use," *Journal of Clinical Psychology(55*), pp. 10-12.

Tanaka, F., Cicourel, A., & Movellan, J. R. 2007. "Socialization between toddlers and robots at an early childhood education center," *Proceedings of the National Academy of Sciences (104*:46), pp.17954-17958.

Wada, K., Shibata, T., Saito, T., & Tanie, K. 2004. "Effects of robot-assisted activity for elderly people and nurses at a day service center," *Proceedings of the IEEE (92*:11), pp.1780-1788.

Weizenbaum, J. 1966. "ELIZA—a computer program for the study of natural language communication between man and machine," *Communications of the ACM (9*:1), pp. 36-45.

Widiger, T. A. 2009. "Neuroticism". In M. R. Leary & R. H. Hoyle (Eds.), Handbook of individual differences in social behavior (p. 129–146). The Guilford Press.

Zhang, Y., Song,W., Tan,Z., Zhu, H., Wang,Y., Lam,CK., Weng,Y, Hoi,SP., Lu,H., Chan,BSM., Chen,J, Yi,L. 2019. "Could social robots facilitate children with autism spectrum disorders in learning distrust and deception?" Computers in Human Behavior (98), pp.140-149.