The Racial Mirroring Effects on Human-Agent in Psychotherapeutic Conversation

Yuting Liao†
College of Information Studies
University of Maryland
College Park, MD, USA
yliao598@umd.edu

Jiangen He
Department of Information Science
Drexel University
Philadelphia, PA, USA
jiangen.he@drexel.edu

ABSTRACT

Conversational agents are increasingly utilized to deliver mental health services. Although these intelligent systems promise more affordable mental healthcare, one critical challenge is patient adherence. Our study explores the "racial mirroring" effects on people's engagement with and perception of agents in the context of psychotherapy. We developed a conversational system with racially heterogeneous personas using strong visual cues. We conducted an experiment by randomly assigning participants (N=212) to racial mirroring, non-mirroring and control groups. Our results suggest that racial mirroring did influence people's perceived interpersonal closeness with agents, use satisfaction, disclosure comfort, desire to continue interacting, and projected future relationship. In this paper, we present the conversational system, experimental procedure and results. We conclude with design recommendations for employing conversational agents in mental health intervention.

CCS CONCEPTS

• Human-centered computing • Human-Computer Interaction
• Computing methodologies • Intelligent agents;

KEYWORDS

Psychotherapy, Mental health support, Race matching, Racial mirroring, Conversational agent, Chatbot

ACM Reference format:

FirstName Surname, FirstName Surname and FirstName Surname. 2020. Insert Your Title Here: Insert Subtitle Here. In *Proceedings of the 25th International Conference on Intelligent User Interfaces(IUI'20). ACM, Cagliari, Italy.* https://doi.org/10.1145/1234567890

1 INTRODUCTION

IUI'20, March, 2020 Cagliari, Italy

https://doi.org/10.1145/1234567890

Many people with mental health issues face significant challenges getting the help they need. Mental health service falls short in accessibility and affordability due to a wide gap between needs and provision. While 1 out of 10 people have a mental health problem, only 1% of the global health workforce provides mental healthcare [67]. Psychological counseling or psychiatry services could be a luxury expense for people under financial stress [28]. Beyond structural barriers, fear of being stigmatized also prevents people from seeking psychological help [33]. To expand the access to mental health services and to counteract the problems of stigma, there has been a burgeoning growth in internet-based and mobile applications for mental health interventions. However, these digital interventions are characterized by relatively poor adoption and adherence [20], which may be due to the lack of the quality of human interaction that a therapist-patient relationship offers [22].

More recently, text-based conversational agents, or chatbots, have gained traction as the new generation of digital mental health support system. Some prominent examples emerging from industry and academia include Woebot (woebot.io) and Wysa (wysa.io). Powered by artificial intelligence (AI) and natural language processing techniques, these conversational agent offering a more natural way of interaction. As people engage in dialogues, chatbots process all text and emoji that a user might enter, and offer responsive, guided conversations and advice to help users cope with challenges to mental health [32]. This human-agent interaction sets to invoke anthropomorphism, making people feel like being in a conversation with humans via a messenger app. The typical mental health services these chatbots provide include targeted therapy exercises, including reframing one's thoughts, mindful breathing, and motivational interviewing.

Literature specifically related to psychotherapeutic chatbots is rather sparse. Recent researchers began to evaluate the efficacy of using conversational agents for mental health outcome (e.g., [22,44]). However, this line of research has not thoroughly interrogated how specific design features might influence client engagement and perceptions of the system. On the other hand, HCI scholarship suggests that conversational agents need to adopt the characteristics of human-human interaction in order to be more engaging [21,53,68]. So far, researchers have examined linguistic and conversation styles,

such as empathic and emotional expression [19], self-disclosure [34], humor [39], and how these characteristics might influence human-agent relationship [9,10,56]. Studies also explore visual characteristics (e.g., [5,6,50]. But to the best of our knowledge, there is no empirical study on how the perceived racial (dis)similarity might influence people's engagement with and perceptions of agents for psychotherapeutic purposes.

Nonetheless, racial identity is an integral part of anthropomorphized agents. It also shapes human's social experiences. In traditional in-person setting, matching clients with therapists of the same race has been found to result in stronger bonding and more positive attitudes [13]. Does perceived racial similarity still matter in agent-based psychotherapy context? If so, understanding the complex relationship between identity and influence will inform a better, human-centric experience.

To evaluate the racial mirroring effects, we developed a conversational system with racially heterogeneous personas. Distinct profile pictures were used as strong visual cues to indicate agents' racial identities, including White/Caucasian, Black/African American, Latinx, and Asian. Due to mixed evidence of cross-gender effects on chatbot perceptions and following [56], we created both female and male personas for each racial group. Beside the differences in agent racial personas, the agents interact with users in the same flow of conversation structured in accordance with motivational interviewing (MI) guidelines. This framework offers a collaborative conversation for strengthening a person's own motivation and commitment to change. It is broadly applicable and often used in cognitive behavioral therapy for mental health issues, such as anxiety and depression and substance use [1,17,21].

Using these prototypes, we conducted an online experiment to investigate the effects of racial mirroring, defined here as the match between the user and agent race/ethnicity. Participants were randomly assigned to racial-mirroring, non-mirroring or control groups. After interacting with the agent, participants completed a survey assessing their perceptions and evaluations of the agent. We investigate the following three research questions:

- **RQ1.** How are people's perceptions of the agent influenced by racial mirroring in a psychotherapeutic conversation?
- **RQ2.** How is people's continued engagement with the agent influenced by racial mirroring?
- **RQ3.** What are people's preferences for racial persona when they are given a chance to customize the conversational interface?

Our analyses revealed that racial mirroring had a positive influence on people's perceived interpersonal closeness with the agent. Although the presence of same-race agents decreased people's comfort with self-disclosure, they reported a higher level of satisfaction associated with their use of chatbot for managing mental well-being. With regard to future engagement, we found participants reported a higher desire to continue interacting with the agent. Participants also predicted a closer future relationship with the same-race agents. Finally, people were significantly more likely to select same-race agent personas when they were given an opportunity to customize the conversational interface.

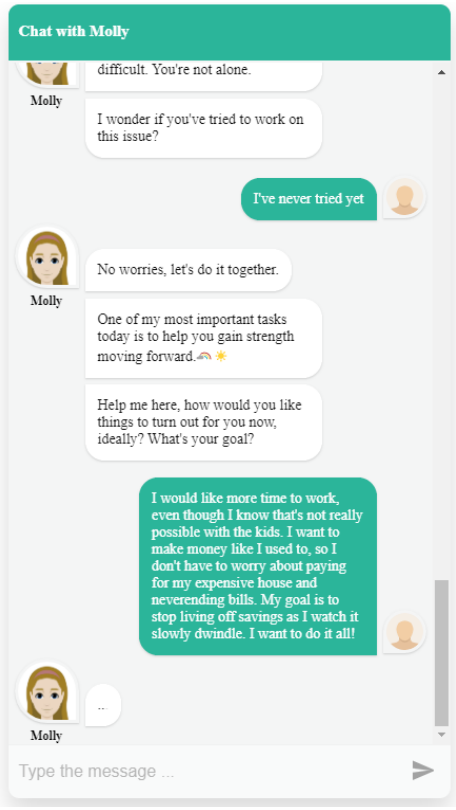


Figure 1: Screenshot of participant conversation with agent Molly in racial-mirroring condition.

The current study offers three major contributions. First, we identify a racially-conscious approach to increasing people's rapport and engagement with therapeutic chatbots. Second, we provide empirical evidence on how racial-mirroring could facilitate people adherence to psychotherapeutic agents. Finally, we open up a discussion about the important issue of race and diversity in intelligent systems *interface* design.

2 RELATED WORK

2.1 Race Matching in Traditional Psychotherapy

Psychologists have explored the effects of matching clients with therapists of the same race/ethnicity across several decades [13]. One stream of research focuses on how race matching influences the counseling process. Studies show that clients or patients could better express themselves when they were paired with therapists of similar race or ethnicity [2]. Based on Cabral and Smith's meta-analysis [13], people prefer to have a therapist of their own race or ethnicity and tend to perceive therapists of

their own race or ethnicity somewhat more positively than others. It is generally safe to assume that matching clients with therapists of the same race should result in stronger therapeutic alliances, or agreement and bonding [25,48,57].

Another stream of research emphasizes how race/ethnicity matching increases the perceived similarities in the values of client and therapist, which in turn predict positive treatment outcomes [29]. Specifically, racial/ethnic matching may improve client outcomes by enhancing mutual understanding between client and therapist and by reducing client concerns about being misunderstood or mistreated [31]. Some social psychology theories might explain client preferences for same-race therapists or counselors. For example, the similarity is associated with credibility [54]. Interpersonal similarity not only influences individual preferences for social interactions but also their perceptions of those interactions.

2.2 Different Sensitivity to Race Matching Effect

The magnitudes of racial mirroring effects are not homogeneous across all race/ethnicity groups. The relevance of racial/ethnic matching was greatest among African American participants [13]. Their preferences for therapists of their own race/ethnicity were very strong; their perceptions of therapists varied substantially as a function of racial/ethnic matching; and their outcomes in therapy tended to be mildly better when their therapist was African American. Grantham [24] interviewed 37 male and female black college students and found that participants preferred black counselors to white counselors to a significantly greater degree; they also had more self-disclosure and self-exploration. In contrast, racial/ethnic matching was apparently least relevant to White/Caucasian Americans; none of the three effect sizes statistically differed from zero. The results with Asian American and Hispanic/Latinx Americans were mixed. Asian Americans showed only mild preferences for a therapist of their own race/ethnicity but tended to provide more positive evaluations for therapists of their own race/ethnicity.

Earlier research has observed an inter-racial tension. Studies [43,55,58,59] consistently shown that African American clients tend to mistrust mental health services provided by white American therapists Explanations for this finding include the possibility of a perceived racial bias in the provision of mental health services and the implicit association of mental health services with the values of White/Caucasian Americans [47].

2.3 Promises and Limitations of Agent-based Mental Health Intervention

These conversational agents promise a feasible, engaging, and effective way for psychoeducation and psychotherapy [60]. Several studies [22,26,27,44] suggest positive outcomes of using chatbots in mental health intervention. For instance, Fitzpatrick et al. [22] conducted a randomized control study with sampled US college students interacting with Woebot for two weeks. They found a reduction in depression symptoms in comparison to the control group who only received information about depression via an e-book. In terms of the positive features, participants particularly appreciated daily check-ins, the bots' empathic and caring 'personality,' and the learning of psychological concepts. In addition to treatment efficacy, there was also a higher engagement compared with other web-based interventions, with eighty-five percent of participants using the bot daily or almost daily.

However, current chatbots are still not capable of delivering full-stack psychological counseling or providing responses comparable to a mental health professional. Due to technical limitations, chatbots often cannot handle longer or more complex messages [22,60]. Additionally, the system allows a very minimal level of personalization exists and provided mental health support is still generic [32].

2.4 Anthropomorphism of Psychotherapeutic Chatbots

Many benefits of using psychotherapeutic chatbots are associated with the anthropomorphism of the technology. The Computers as Social Actors theorem by Reeves and Nass [46] suggest that individuals apply social rules to interactions with computers. In the study of Woebot [22], some participants reported that the best thing about their experience with the therapist chatbot was the perceived empathy. However, anthropomorphism can have negative impacts. Norman [42] cautions that If an interface is anthropomorphized too realistically, people tend to form unrealistic expectations. Norman suggests that people will be more accepting of an intelligent interface when their expectation matches with its real functionality. Despite the tension between anthropomorphism and user satisfaction, chatbots are often designed with anthropomorphized characteristics, including gender, age, and ethnicity identities [17]. Previous researchers also explore a variety of social traits, including conversational intelligence (e.g., greeting [1]) and social intelligence (e.g., managing conflicts [51]] and empathy[35]) and personality [40]. Due to the nature of text-based interface, language style and pattern remains the primary focus in the provision of social characteristics for. For instance, Araujo [1] found that people rated the agent to be more likable and friendly when agents had a human name, used informal language and greetings. Although these linguistic cues have important implications, our focus is on the visual cues that indicate different racial identities of agents.

2.5 Race and Chatbots

With an emphasis on the issue of race, a line of research concerns about how chatbots handle conversations involving race stereotypes [36,52]. Schlesinger et al. revealed the race and bias issues in the machine learning algorithm underlying the chatbot and called for a better understanding of the racial context of chatbots. From an ethical ad normative perspective, these studies are pivotal in raising awareness around the issue of race in conversation design. An equally important topic is anthropomorphized agents' racial identity. Brahnam et al. [12]

showed how an agent's racial identity might have the unintended consequence of problematic use. They found that when the avatars were presented as black adults, references to race can deteriorate into racist attacks. With a particular focus on the effect of racial mirroring, Baylor and Kim (2003b) [6] examined the impact of pedagogical agents' ethnicity on learners' perception of the agents. Through an experiment with undergraduate participants, they found that students who worked with pedagogical agents of the same ethnicity rated the agents as more credible, engaging, and affable than those who worked with agents of different ethnicity. The issue of race is important yet complex. There is a need to place racially-conscious intelligent interface design on the agenda.

3 SYSTEM OVERVIEW

3.1 Chatbot Design

First, we created the chatbot system on Google's DialogFlow, a Natural Language Understanding platform for building conversational applications. Next, we integrated the application with an external web-based interface. Specifically, we used a Node.js client of DialogFlow, which is connected by JavaScript in the interface in order to replay responses from DialogFlow to users.

The interface was built as a web application using Javascript, React.js, and HTML5. The flexibility of such a web-based interface allowed us to alter the personas of agents for our experiment. We purposefully limited visual elements that are not relevant to our study (e.g., downsizing the user profile icon).

3.2 Agent's Racial Identity Cues

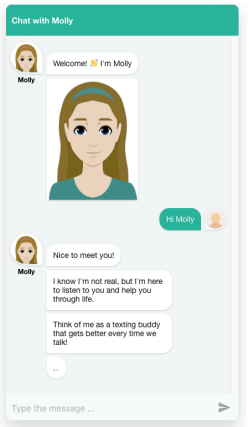
The primary cue for the agents' racial identity is their profile pictures. An enlarged depiction of this profile picture was presented to the users shortly after greeting during the Engaging process. Throughout the conversation, this profile picture was displayed alongside each conversational message of the agent in a familiar text message environment.

The secondary cue is the agent's name. We assigned agents with names that might have implied their racial identity [23,41], for example, Jake to suggest a Caucasian identity, Ebony to suggest an African-American identity, and Antonio to suggest a Latinx identity. These name choices were based on previous study example [23] and report [41] of name popularity based on race and ethnicity.

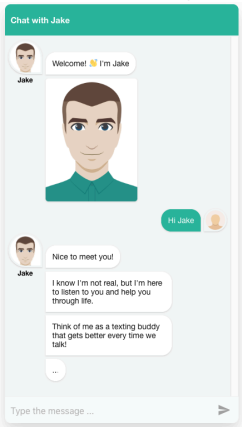
In total, we created eight racially diverse personas, including Molly (White/Caucasian female), Jake (White/Caucasian male), Ebony (Black/African American female), Darnell (Black/African American male), Jenny (Asian female), David (Asian male), Camila (Hispanic female), and Antonio (Hispanic male).

Additionally, we created Robbie, a non-personified and gender-neutral agent that assembles a typical robot. Robbie interacted with participants in the control group.

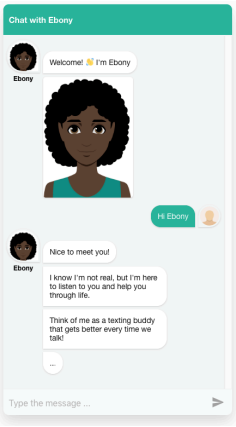
Molly
(White female agent)



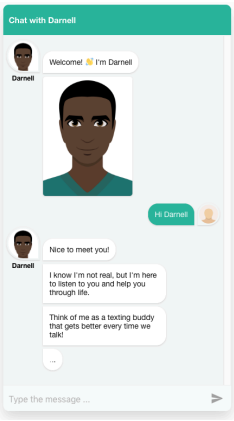
Jake
(White male agent)



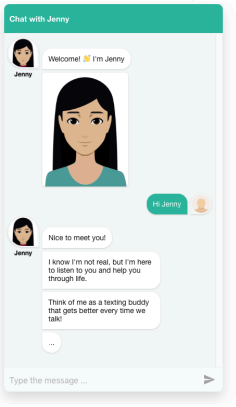
Ebony
(Black female agent)



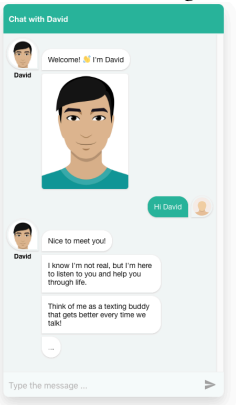
Darnell
(White male agent)

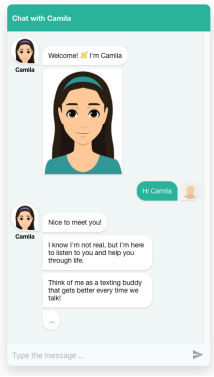
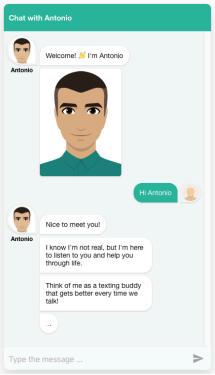


Jenny
(Asian female agent)



David
(Asian male agent)



Camila (Hispanic female agent)	Antonio (Hispanic male agent)
	

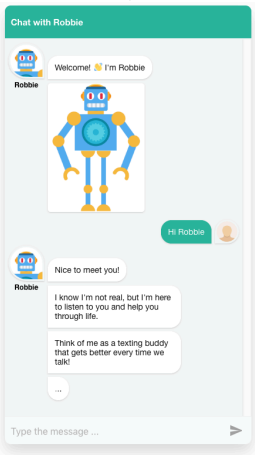
Robbie (Non-personified, gender neutral agent)


Table 1. Visual Cues Indicating Agent Racial Identity

3.3 Dialogue Design

3.3.1 Four Motivational Interviewing (MI) Processes. Our system was built to lead a structured conversation following the typical MI flow in four processes [37]: Engaging, Focusing, Evoking, and Planning. Although Google's DialogFlow allows the chatbot to provide various questions or responses, we placed a higher priority on the consistency of experimental conditions. Therefore, each participant was asked the same set of questions and provided predefined feedback. Participants engaged in the interaction by typing in their responses via the text entry. We designed conversation flow closely align with the four MI processes. In Engaging, the agent greets the user and offers a self-introduction to build a relational foundation. Next at Focusing stage, the agent asks the user to identify and detail their biggest challenge in life, with the follow-up prompt probing their feelings and possibly an inner struggle. This leads to Evoking, where the agent explores future goals with the user, inspiring their own ideas for change. Finally, at Planning stage, the agent invites the user to specify a tangible change that can be made to overcome the challenge.

MI Process	Sample Dialogue
Engaging	"My goal today is to inspire any changes you'd need to overcome the big challenge."
Focusing	"What happened? How has it influenced your life?"; "How do you feel about all this?"
Evoking	"How would you like things to turn out for you now, ideally? What's your goal?"; "Let's talk about your strength. What strong points do you have that could help you move forward?"
Planning	"What would you say is the most important thing that you can do to achieve this goal?"

Table 1. Sample Dialogues in Four MI processes

3.3.2 Linguistic Style. Since the chatbot we developed adopts an important psychotherapeutic role, building rapport is essential. The linguistic style is a key aspect for the creation of artificial agents [63]. We adapted linguistic style from prior work [10,11,26,38,39,64,65] to develop attachment bonds between the agents and users.

Linguistic Characteristics	Sample Dialogue
Greetings [15]	"Welcome! I'm Molly"; "Nice to meet you!"
Emoticons [64]	"One of my most important tasks today is to help you regain strength.🌈☀️"
Small talk [7]	"I know I'm not real, but I'm here to listen to you and help you through life."
Self-disclosure [38]	"As a conversational AI, I can engage you in a "talk" session, just like that in a counseling or psychotherapy! "
Empathy [11,26]	"I understand things could be difficult. You're not alone."
Meta-relational talk [8,9])	"Think of me as a texting buddy that gets better every time we talk!"
Humor [39]	"... can you name one specific thing that you could do to overcome the challenge, maybe something small, like an apple a day?"

Table 3. Sample Dialogues for Linguistic Characteristics

3.3.3 Tailored Responses. In a professional setting, therapists will need to gather specific information about their personal experience and/or mental state in order to provide tailored intervention to clients. To this end, the chatbot offers the "quick reply" feature. Several predefined options were displayed so participants could provide just-in-time self-evaluation. The agent would follow up with tailored responses based on the users' experience or psychological state. For instance, when a user indicated that he has tried to fix the issue, but has since abandoned action, the agent would follow up with the following response: "One of my most important tasks today is to help you regain strength." When the user indicated that he had never tried, the agent would say "No worries, let's do it together." Likewise, when the user indicated that he was not confident at all about dealing with the challenge, the agent would say "I have a lot of confidence that working together we can find a way out."; for users with high confidence, the agent would respond, "It seems like you have a lot of courage! That's a good sign."

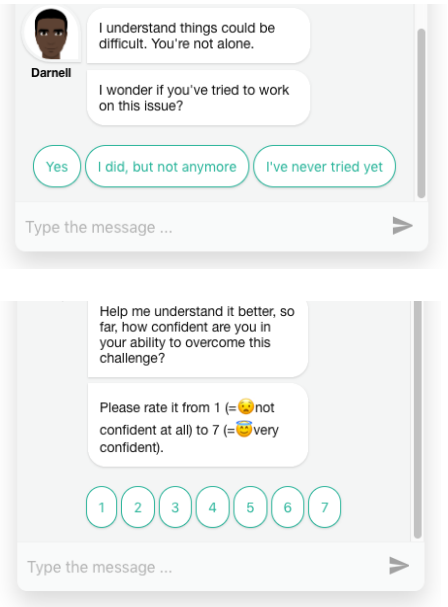


Figure 2: Two examples of predefined options for providing tailored responses.

4 EXPERIMENT PROCEDURE

4.1 Participant Recruitment and Screening

We conducted an online experiment. Participants were recruited from the Amazon Mechanical Turk platform, each participant who completed the entire process was compensated with $2. Before proceeding, all workers consented to participate and confirmed that they were older than 18 years old living in the United States. Participants first filled out a screening questionnaire that asked them about their psychological well-being and racial background. Since the experimental conditions predicated on the perceived racial (dis)similarity, our research team followed two selection criteria in the recruitment process. First, it was necessary that our research only recruited people who have a clear, unambiguous self-identification with one particular race. Therefore, we excluded people who identified themselves as multi-racial. Second, to ensure the validity of the experimental conditions based on perceived racial similarity, participants were selected whose facial features assemble that of agents. As such, the survey screened out respondents identified as Native Hawaiian or other Pacific Islander, Native American or Alaska Native. Additionally, since Asian are a greatly diverse group, for people who identified as Asian, the questionnaire also asked which region they would trace their cultural heritage back. We only recruited Eastern Asian and mainland Southeast Asian participants because the facial features of Asian agents David and Jenny are more aligned with these two groups.
A total of 212 participants were recruited, including 155 White/Caucasian, 29 Black/African American, 13 Latinx, and 15 Asian.

4.2 Experimental Groups

After the screening survey, eligible participants were randomly assigned to one of the three experimental groups:

- *Racial-mirroring group:* A participant would interact with one of the two same-race agents. For example, a Latinx participant would talk to Camila or Antonio. Under this condition, the agent's gender was randomized and evenly distributed.
- *Non-mirroring group:* A participant would interact with one of the six different-race agents. for example, a Latinx participant would talk to Molly, Jake, Ebony, Darnell, Jenny or David with equal probability.
- *Control group:* A participant would interact with Robbie, the non-personified, gender-neutral agent.

At the close of our data collection, 71 participants were assigned to the racial-mirroring group, 71 were assigned to non-mirroring groups, and 70 participants were assigned to the control group.

4.3 Interaction with the Agent

Participants interacted with the agent in MI processes described above. The self-disclosure primarily occurred in the last three processes, including Focusing, Evoking and Planning. In Focusing, the agent asked them to talk about their biggest challenges in life. Participants brought up a variety of life struggles. These include behavioral issues (e.g., P93: "*I struggle with gambling problems*"), financial stress (e.g., P52: "*There never seems to be quite enough money for everything*"), job-related (e.g., P118:*"I completed my degree but I'm not able to find a suitable job for 6 months"*), health problems (e.g., P144: *"I suffer from severe vertigo which limits my life in terms of movement*"), relationship issues (e.g., P142: "*My so-called boyfriend cheated.*"), and negative thoughts (e.g., P93: "*I feel like a failure*"). Participants' self-disclosure depth varied. Some participants provided a detailed picture of their struggle, as P85 wrote, "*My husband has had prostate surgery and is now on bed rest for a full month and I need*

to take care of him and also work."; Other participants submitted very cursory responses, such as "*I have jealousy issues* (P48)"; Some responses were candid and upfront, such as "*I have depression. I have been depressed for as long as I can remember. I have lost jobs and friends. I truly hate myself and hate the world.*" Next, the agent asked how they felt about their major challenge in life. Similar to self-disclosure related to their challenges, our participants expressed their feelings in varying length and depth, with some were more specific (e.g., P79: "*I feel stupid to be addicted to cigarettes, it's a brain poison.*") than the others (e.g., P79: "I feel frustrated and stressed"). In Focusing stage, client offered a variety of targeted behavior or goal depending their challenges. In Evoking, participants were guided by the agent to explore the potential changes that they need to achieve the goal. Similar to that in the Focusing stage, participants' disclosure in these two stages varied in terms of length, depth and topic sensitivity.

4.4 Post-Interaction Survey

At the end of the conversation with the agent, participants were provided a link to the post-interaction survey evaluating their perceptions of the agent. We applied the piped text feature in Qualtrics to customize the agent name as per the specific agent participants interacted with. The main measures for dependent variables (DVs) are described below:

- *Perceived Interpersonal Closeness.* (M=3.38, SD=1.23). The Scale of Perceived Interpersonal Closeness Scale (PICS), a single-item pictorial scale [45], is an established and reliable instrument to measure the perceived closeness of a relationship. Participants chose an area in the picture to reflect the relationship between the agent and the self, with a range from 1 (Distant) to 7 (Fully close).
- *Comfort with Self-Disclosure. (*M=3.87, SD=.62, α=.87). This original 5-item measure evaluated how comfortable participants felt about disclosing to the agent. On a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree), participants rated their level of agreement with statements such as: I felt comfortable when I talked with [Agent Name] about my challenge and life experiences; I expressed myself openly; I felt [Agent Name]wouldn't judge me when I talked about my issue.
- *User Satisfaction.* (M=3.44, SD=.97, α=.94). This is a 6-item measure to assess the experience with the chatbot. We adapted the instrument of Venkatesh et al. [61] to reflect the chatbot use and satisfaction in the psychotherapeutic context. Sample statements include: My interaction with [Agent Name] was satisfying; After my conversation with [Agent Name], I felt more confident about my ability to overcome the challenge; This technology is useful for me to manage my mental well-being. Participants chose their level of agreement with these statements from 1 (Strongly disagree) to 5 (Strongly agree).
- *Recommendation.* (M=60.06, SD=13.40, α=.93). This 2-item measure was included to assess how much participants would recommend the agent 1) to their friends or family for managing mental well-being and 2) people who have a need for counseling/psychotherapy. Participants chose from 1 (Would not recommend at all) to 100 (Would definitely recommend).
- *Desire to Continue Interacting with the Agent.* (M=63.57, SD=28.37). We adapted the instrument of Bickmore et al. study [10] with an emphasis on interacting with the agent for psychotherapeutic purposes. Participants rated how much they would like to continue working with the agent from 1(No at all) to 100 (Very much willing).
- *Projected Future Relationship.* (M=3.57, SD=1.80). This is measured with the inclusion-of-the-Other-in-the-Self (IOS) Scale[3], an established measure for relationship strength. Participants were asked to predict how close their relationship with the agent would be, under the assumption that they would have conversations at least three times per week.
- *User Preference of the Agent's Racial Persona.* At the end of the survey, all nine agent personas were presented. Participants were asked to identify which agent(s) they would prefer to interact with for mental well-being support. This choice was included in the Chi-square test of independence to answer RQ3.

Several measures evaluating individual background, attitudes and experience were included as independent variables and control variables.

- *Anthropomorphism, Liveliness, Likeability, and Perceived Intelligence.* The Godspeed instrument [4] has been used in the HRI research to measure various perceptions of AI. Our survey includes dimensions of Anthropomorphism (M=59.58, SD=25.60, α=.96), Liveliness (M=66.78, SD=20.62, α=.92), Likeability (M=77.35, SD=14.96, α=.90), and Perceived Intelligence (M=72.43, SD=18.34, α=.95). Participants rated their impression of the agent along with linguistic scales in word pairs, such as fake--natural, machinelike--humanlike, unfriendly--friendly, awful--terrible. Responses were coded from 1 to 100.
- *General Attitudes toward AI.* (M=60.06, SD=13.40, α=.93). We adapted the instrument from Cave et al. [16] study to evaluate participant's general attitudes toward artificial intelligence.
- *Working Alliance:* To assess attachment bond, we adopt the bond subscale of the Working Alliance Inventory for Technology-Based Interventions [30].
- *Control variables.* Our participants provided basic demographic background, including their gender, current age, annual income, and education. Privacy concerns [62], psychological well-being [49] and satisfaction with life [18] measures were also included as control variables.

5 RESULT

We conducted a multivariate analysis of covariance (MANCOVA) using perceived interpersonal closeness, disclosure comfort, desire to continue interacting with agent, satisfaction, and recommendation as dependent variables, while controlling for agent and participant genders, privacy concern, general attitudes toward AI, the anthropomorphism of agents, and background psychological factors. The results revealed significant main effects for racial mirroring, Wilk's $\lambda = .83$, $F(6, 218) = 3.98$, $p < .001$. Since MANCOVA is an omnibus test, we further proceeded with univariate analysis ANCOVA to unpack how racial mirroring influence people's perceptions of the agent

they interacted in psychotherapeutic conversation, and how these effects varied based on agent gender and client gender. To ensure the validity of our analyses, we chose to perform generalized linear modeling (GLM) rather than linear regression due to the non-normal distribution of our dependent variables [69].

RQ1: How Racial Mirroring Influenced Perceptions of Agents

5.1.1 The effect of racial mirroring on perceived interpersonal closeness. The modeling showed a significant main effect for racial mirroring on participants' perceived interpersonal closeness with agent, F (2, 216) = 35.61, $p < .001$, partial $\eta^2 = .15$. Specifically, Bonferroni post-hoc comparisons revealed that participants perceived a closer personal relationship with same-race agents (*M=3.47*, *SE=*.09), compared to those who interacted with different-race agents (*M=3.22*, *SE=*.09).

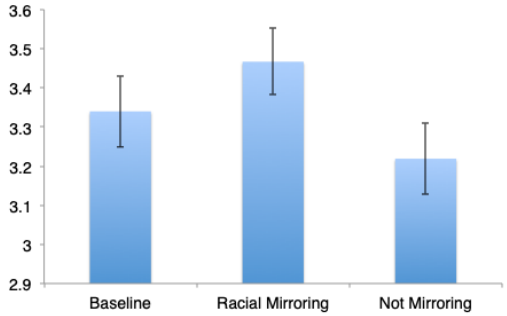


Figure 3: Different Levels of Perceived Interpersonal Closeness

5.1.2 The effect of racial mirroring on client self-disclosure comfort. The modeling showed a significant main effect for racial mirroring on participants' comfort with self-disclosure, F (2, 216) = 40.63, $p < .001$, partial $\eta^2 = .18$. Specifically, Bonferroni post-hoc comparisons revealed that participants felt less comfortable disclosing to same-race agents (*M=3.78*, *SE=*.04), compared to different-race agents (*M=3.93*, *SE=*.05).

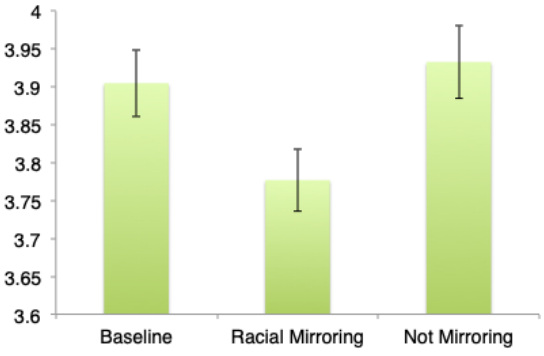


Figure 4: Different Levels of Disclosure Comfort

5.1.3 The effect of racial mirroring on User Satisfaction. We focus on the dependent variable of user satisfaction. The modeling showed a significant main effect for racial mirroring on how satisfied participants felt about their interaction with the agent, F (2, 216) = 34.22, $p < .001$, partial $\eta^2 = .16$. Specifically, Bonferroni post-hoc comparisons revealed that participants were more willing to recommend same-race agents to others (M=3.48, SE=.04), compared to different-race agents (M=3.30, SE=.05). We observed that participants in the control group reported a higher level of satisfaction than people in the racial mirroring group. This reduced in satisfaction might be linked to participants' higher expectation of more human-like agents.

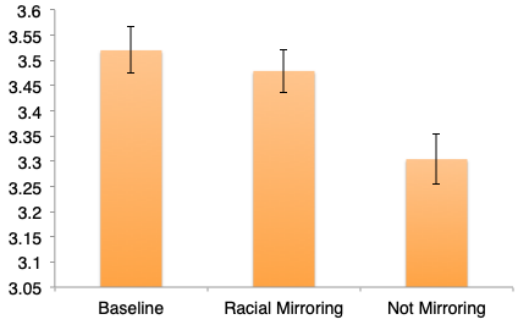


Figure 5: Different Levels of User Satisfaction

5.1.4 The effect of racial mirroring on recommendation likelihood. The modeling showed a significant main effect for racial mirroring on how much participants would recommend the agent to their friends and family, and people in need of psychotherapy, F (2, 216) = 54.29, $p < .001$, partial $\eta^2 = .32$. Specifically, Bonferroni post-hoc comparisons revealed that participants were more willing to recommend same-race agents to others (*M=61.68*, *SE=*1.27), compared to different-race agents (*M=55.61*, *SE=1.*49).

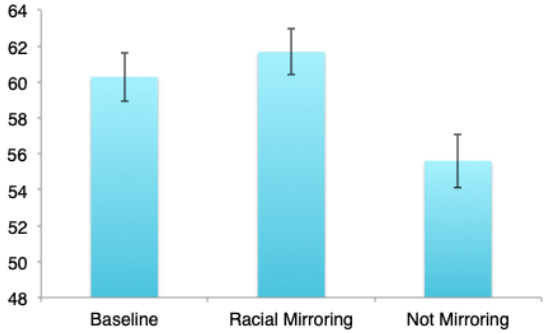


Figure 6: Different Levels of Recommendation Likelihood

RQ2: How Racial Mirroring Influenced Future Engagement

5.2.1 The effect of racial mirroring on the Desire to Continue Interacting with Agent. The modeling showed a significant main effect for racial mirroring on participants' desire to continue interacting with agent, F (2, 216) = 39.57, $p < .001$, partial $\eta^2 = .14$. Specifically, Bonferroni post-hoc comparisons revealed that

participants have a stronger desire to continue interacting with same-race agents (*M=64.93, SE=*1.25), compared to different-race agents (*M=61.47, SE=1.*47).

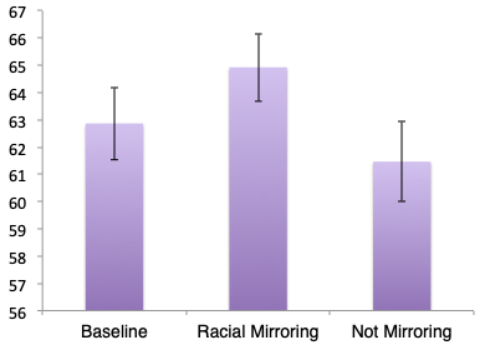


Figure 7: Different Levels of Desires to Continue Interacting with Agent

5.2.2 The effect of racial mirroring on projected future relationship. The modeling showed a significant main effect for racial mirroring on participants' prediction of future relationship with agent, F (2, 216) = 42.16, $p < .001$, partial $\eta^2 = .18$. Specifically, Bonferroni post-hoc comparisons revealed that participants predicted that they would have a closer relationship with same-race agents (*M=3.71, SE=*.07), compared to different-race agents (*M=3.21, SE=*.09)

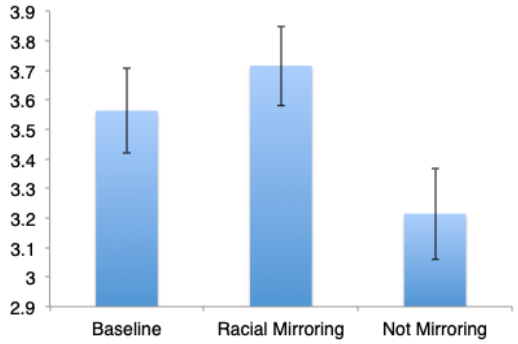


Figure 8: Different Levels of Projected Future Relationship Closeness

5.3 RQ3: User Preferences of Agent Personas

At the end of the survey, participants were asked to select agent personas they preferred if they could customize the therapeutic chatbot profile. We conducted a Chi-square test of independence based on participant race and the agent personas they selected. The result revealed that participants' preferences of agent personas were statistically significant based on their own races, $\chi 2$ (24,206) =142.23, p<.001. The further residual analyses surfaced that participants preferred to use same-race chatbots. Specifically, White/Caucasian participants were much more likely to select Molly and Jake; Black/African American participants were much more likely to select Ebony and Darnell; Latinx participants were much more likely to select Molly and Jake; and Asian participants were much more likely to select Jenny and David. Such preference of same-race agents was especially pronounced among Black/African American participants. The chance that they selected Ebony, the female Black/African agent, was seven times than expected.

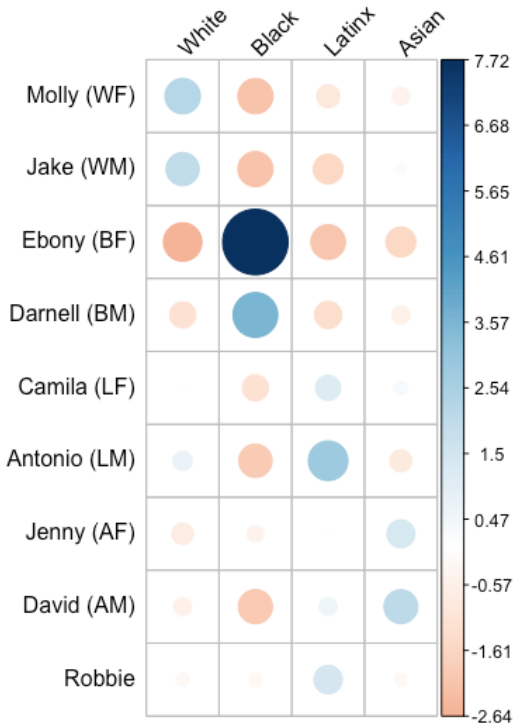


Figure 9: A visualization of standardized Chi-square residuals, Blue nodes indicate higher preferences; red nodes indicate lower preferences.

6 DISCUSSION

For personified agents, race is an integral part of the design. In the context of psychotherapy, our study suggests that racial preferences do not disappear in human-agent interaction. Our analyses revealed that people who interacted with same-race agents would have perceived a closer interpersonal relationship with the agent. Additionally, they had a higher level of user satisfaction and more desire to continue interacting with the agent and were also more likely to recommend the agent to their friends and family and people who need psychological help. These findings are in line with studies conducted in traditional in-person psychotherapy setting. Racial mirroring is particularly effective for African American clients. Our study reveals that African American participants had a much stronger preference for same-race agents than any other group. Given differences in the experiences of racial (in)equality, it should not be surprising that racial mirroring may not impact all racial or ethnic groups in the same manner.

Nonetheless, racial mirroring between client and agent appeared to entail a decrease in self-disclosure comfort. Participants who interacted with same-race agents reported a lower level of disclosure comfort; they were more likely to feel being judged than participants in the non-mirroring and control groups. The reduced disclosure comfort might result from

people's stronger need for impression management (Leary, 2019) and heightened fear of public stigma (i.e., negative views of the person by others) [Vogel, 2007]. In particular, racial similarity might associate with a greater degree of "groupness", the degree to which a collection of people is perceived as a unified or meaningful entity (Campbell, 1958; Hamilton & Sherman, 1996). And people might encounter a higher pressure for conformity and worry about being labeled as "abnormalcy".

Design Implications

The racial-mirroring effects present an opportunity for employing chatbots to deliver mental health support. In the traditional in-person counseling, the practice of matching clients with therapists of the same race is not always feasible, because there are significantly fewer therapists of color than would be necessary to systematically implement racial/ethnic matching (APA, 2005). In contrast, implementing racial mirroring on digital platforms would be more practical with less friction. Our study suggests that racial mirroring facilitates the interpersonal relationship between clients and agents. As the bonding between client and therapist plays a critical role in promoting patient adherence and improving treatment outcome, racial mirroring can be a desirable feature for therapeutic chatbots. On the other hand, non-mirroring should be avoided. Racial minorities, especially African Americans, are more sensitive to the racial mirroring effect. Allowing clients to interact with same-race agents would more substantially benefit these groups who have been marginalized and underserved in traditional psychotherapy.

Nonetheless, it might be presumptuous to state that chatbot designers should match the race/ethnicity of agents to the users for all scenarios. Our study reveals that disclosure comfort decreased when people interacted with same-race agents; they might be more concerned about being judged. This finding suggests that chatbot should provide encouraging words to mitigate users' concerns. Some visual cues, such as emoji or picture could also be incorporated to create a relaxing, friendly environment.

Another concern about racial-mirroring is that it could lead to further fractionating of social groups. Also, our study did not find a statistically significant difference between the control group and racial-mirroring group perceptions in terms of interpersonal closeness, satisfaction, desire to continue using, and recommendation. Therefore, we recommend assigning a non-personified persona for therapeutic chatbots if there were no sufficient information regarding user demographics.

An ideal design outcome could be that users can customize the personas of therapeutic agents. This flexibility serves two purposes. First, users can gain a higher sense of control, which might improve system engagement. As an African American participant wrote in the feedback, "*I really enjoyed talking with Ebony. It would be great if a person would be able to change the avatar and name of the AI, to a figure that is comfortable to talk to.*" Second, this configuration helps to set a realistic expectation: this is a bot, not a human therapist that can provide perfect responses. As Norman cautioned, a too realistic human-like appearance and interaction can be deceptive and misleading by implying promises of functionality that can never be reached [42].

Limitations and Future Direction

We only tested racial mirroring effects with regard to the use of psychotherapeutic chatbots among four racial groups in the United States. The racial dynamics are distinct from other countries. The context of psychotherapeutic conversation might be idiosyncratic. Therefore, the findings might not be generalizable to another nation, race, or conversation context. Our participants recruited from Amazon Mechanical Turk were mainly white; results might bias toward this group. As the use of chatbot in mental health services increase, the user base can be greatly diverse and multicultural. For instance, WHO's STARS project implemented a chatbot in five countries including Jamaica, Nepal, Pakistan, South Africa and West Bank and Gaza Strip. Future researchers should explore racial mirroring effects in other cultural contexts. Race is only one aspect of identity. The larger picture here is identity unification between agents and users. Future researchers should also explore the perceived similarity in other dimensions of identity, such as sexual orientation, faith/spirituality, and physical ability.

7 CONCLUSION

In a broader sense, a racially-conscious intelligent system exemplifies inclusive design that draws on the full range of human diversity. Imagine a future of sophisticated AI psychotherapists where a realistic human-like appearance is expected. Consider a fallacious design where AI therapists appear to be a uniform monocultured race, while patients are from 15 other races and ethnicity? It would seem absurd, perhaps oppressive, but certainly ineffective as the current study suggests. As chatbot entities and other forms of conversational agents become more prevalent, it is critical to integrate the consideration of race at both algorithm and interface levels. Failure to adopt a racially-conscious strategy might prevent the industry–and society at large–from realizing the full potential of AI.

REFERENCES

[1] Theo Araujo. 2018. Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Comput. Hum. Behav.* 85, (August 2018), 183–189.

[2] Hal Arkowitz and Henny A. Westra. 2009. Introduction to the special series on motivational interviewing and psychotherapy. *J. Clin. Psychol.* 65, 11 (November 2009), 1149–1155. DOI:https://doi.org/10.1002/jclp.20640

[3] Arthur Aron and Barbara Fraley. 1999. Relationship Closeness as Including Other in the Self: Cognitive Underpinnings and Measures. *Soc. Cogn.* 17, 2 (June 1999), 140–160. DOI:https://doi.org/10.1521/soco.1999.17.2.140

[4] Christoph Bartneck, Dana Kulic, and Elizabeth Croft. 2008. Measuring the anthropomorphism, animacy, likeability,

perceived intelligence, and perceived safety of robots. *Tech. Rep.* (2008), 8.

[5] Amy Baylor and Yanghee Kim. 2003. The Role of Gender and Ethnicity in Pedagogical Agent Perception. 1503–1506. Retrieved October 8, 2019 from https://www.learntechlib.org/primary/p/12158/

[6] Amy L. Baylor and Yanghee Kim. 2004. Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role. In *Intelligent Tutoring Systems* (Lecture Notes in Computer Science), 592–603.

[7] Timothy Bickmore and Justine Cassell. SOCIAL DIALOGUE WITH EMBODIED CONVERSATIONAL AGENTS. 32.

[8] Timothy Bickmore, Daniel Schulman, and Langxuan Yin. 2010. Maintaining Engagement in Long-term Interventions with Relational Agents. *Appl. Artif. Intell. AAI* 24, 6 (July 2010), 648–666. DOI:https://doi.org/10.1080/08839514.2010.492259

[9] Timothy W. Bickmore, Suzanne E. Mitchell, Brian W. Jack, Michael K. Paasche-Orlow, Laura M. Pfeifer, and Julie ODonnell. 2010. Response to a Relational Agent by Hospital Patients with Depressive Symptoms. *Interact. Comput.* 22, 4 (July 2010), 289–298. DOI:https://doi.org/10.1016/j.intcom.2009.12.001

[10] Timothy W. Bickmore and Rosalind W. Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 293–327. DOI:https://doi.org/10.1145/1067860.1067867

[11] Hana Boukricha and Ipke Wachsmuth. 2011. Empathy-based emotional alignment for a virtual human: A three-step approach. *KI - Künstl. Intell.* 25, 3 (2011). Retrieved October 9, 2019 from https://pub.uni-bielefeld.de/record/2276585

[12] Sheryl Brahnam and Antonella De Angeli. 2012. Gender affordances of conversational agents. *Interact. Comput.* 24, 3 (May 2012), 139–153. DOI:https://doi.org/10.1016/j.intcom.2012.05.001

[13] Raquel R. Cabral and Timothy B. Smith. 2011. Racial/ethnic matching of clients and therapists in mental health services: A meta-analytic review of preferences, perceptions, and outcomes. *J. Couns. Psychol.* 58, 4 (2011), 537–554. DOI:https://doi.org/10.1037/a0025266

[14] Robert R. Carkhuff and Richard Pierce. 1967. Differential effects of therapist race and social class upon patient depth of self-exploration in the initial clinical interview. *J. Consult. Psychol.* 31, 6 (1967), 632–634. DOI:https://doi.org/10.1037/h0025163

[15] Justine Cassell and Timothy Bickmore. 2003. Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents. *User Model. User-Adapt. Interact.* 13, 1 (February 2003), 89–132. DOI:https://doi.org/10.1023/A:1024026532471

[16] Stephen Cave, Kate Coughlan, and Kanta Dihal. 2019. "Scary Robots": Examining Public Responses to AI. In *AIES*. DOI:https://doi.org/10.1145/3306618.3314232

[17] Ana Paula Chaves and Marco Aurelio Gerosa. 2019. How should my chatbot interact? A survey on human-chatbot interaction design. *ArXiv190402743 Cs* (April 2019). Retrieved October 9, 2019 from http://arxiv.org/abs/1904.02743

[18] Ed Diener, Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. 1985. The Satisfaction With Life Scale. *J. Pers. Assess.* 49, 1 (1985), 71–75. DOI:https://doi.org/10.1207/s15327752jpa4901_13

[19] Kohji Dohsaka, Ryota Asai, Ryuichiro Higashinaka, Yasuhiro Minami, and Eisaku Maeda. 2009. Effects of conversational agents on human communication in thoughtevoking multi-party dialogues. In *In Proc. SIGDIAL*, 217–224.

[20] Liesje Donkin, Ian B. Hickie, Helen Christensen, Sharon L. Naismith, Bruce Neal, Nicole L. Cockayne, and Nick Glozier. 2013. Rethinking the Dose-Response Relationship Between Usage and Outcome in an Online Intervention for Depression: Randomized Controlled Trial. *J. Med. Internet Res.* 15, 10 (2013), e231. DOI:https://doi.org/10.2196/jmir.2771

[21] A. C. Elkins, D. C. Derrick, J. K. Burgoon, and J. F. Nunamaker Jr. 2012. Predicting Users' Perceived Trust in Embodied Conversational Agents Using Vocal Dynamics. In *2012 45th Hawaii International Conference on System Sciences*, 579–588. DOI:https://doi.org/10.1109/HICSS.2012.483

[22] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment. Health* 4, 2 (2017), e19. DOI:https://doi.org/10.2196/mental.7785

[23] Roland G Fryer and Steven D Levitt. 2003. *The Causes and Consequences of Distinctively Black Names*. National Bureau of Economic Research. DOI:https://doi.org/10.3386/w9938

[24] Robert J. Grantham. 1973. Effects of counselor sex, race, and language style on black students in initial interviews. *J. Couns. Psychol.* 20, 6 (1973), 553–559. DOI:https://doi.org/10.1037/h0035188

[25] Don K. Harrison. 1975. Race as a counselor-client variable in counseling and psychotherapy: A review of the research. *Couns. Psychol.* 5, 1 (1975), 124–133. DOI:https://doi.org/10.1177/001100007500500130

[26] Becky Inkster, Shubhankar Sarda, and Vinod Subramanian. 2018. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR MHealth UHealth* 6, 11 (2018), e12106. DOI:https://doi.org/10.2196/12106

[27] Takeshi Kamita, Tatsuya Ito, Atsuko Matsumoto, Tsunetsugu Munakata, and Tomoo Inoue. 2019. A Chatbot System for Mental Healthcare Based on SAT Counseling Method. *Mobile Information Systems*. DOI:https://doi.org/10.1155/2019/9517321

[28] Bradley E. Karlin, Michael Duffy, and David H. Gleaves. 2008. Patterns and predictors of mental health service use and mental illness among older and younger adults in the United States. *Psychol. Serv.* 5, 3 (2008), 275–294. DOI:https://doi.org/10.1037/1541-1559.5.3.275

[29] T. A. Kelly and H. H. Strupp. 1992. Patient and therapist values in psychotherapy: perceived changes, assimilation, similarity, and outcome. *J. Consult. Clin. Psychol.* 60, 1 (February 1992), 34–40. DOI:https://doi.org/10.1037//0022-006x.60.1.34

[30] Brian D. Kiluk, Kelly Serafini, Tami Frankforter, Charla Nich, and Kathleen M. Carroll. 2014. Only Connect: The

Working Alliance in Computer-Based Cognitive Behavioral Therapy. *Behav. Res. Ther.* 63, (December 2014), 139–146. DOI:https://doi.org/10.1016/j.brat.2014.10.003

[31] Eric L. Kohatsu, Michael Dulay, Cynthia Lam, William Concepcion, Patricia Perez, Cynthia Lopez, and Jennie Euler. 2000. Using racial identity theory to explore racial mistrust and interracial contact among Asian Americans. *J. Couns. Dev.* 78, 3 (2000), 334–342. DOI:https://doi.org/10.1002/j.1556-6676.2000.tb01915.x

[32] Kira Kretzschmar, Holly Tyroll, Gabriela Pavarini, Arianna Manzini, and Ilina Singh. 2019. Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support. *Biomed. Inform. Insights* 11, (January 2019), 1178222619829083. DOI:https://doi.org/10.1177/1178222619829083

[33] Daniel G. Lannin, Max Guyll, David L. Vogel, and Stephanie Madon. 2013. Reducing the stigma associated with seeking psychotherapy through self-affirmation. *J. Couns. Psychol.* 60, 4 (October 2013), 508–519. DOI:https://doi.org/10.1037/a0033789

[34] SeoYoung Lee and Junho Choi. 2017. Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *Int. J. Hum.-Comput. Stud.* 103, (July 2017), 95–105. DOI:https://doi.org/10.1016/j.ijhcs.2017.02.005

[35] Bingjie Liu and S. Shyam Sundar. 2018. Should Machines Express Sympathy and Empathy? Experiments with a Health Advice Chatbot. *Cyberpsychology Behav. Soc. Netw.* 21, 10 (October 2018), 625–636. DOI:https://doi.org/10.1089/cyber.2018.0110

[36] Mark C. Marino. 2014. The Racial Formation of Chatbots. DOI:https://doi.org/10.7771/1481-4374.2560

[37] William R. Miller and Stephen Rollnick. 2013. *Motivational interviewing: Helping people change, 3rd edition.* Guilford Press, New York, NY, US.

[38] Youngme Moon. 2000. Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers. *J. Consum. Res.* 26, 4 (March 2000), 323–339. DOI:https://doi.org/10.1086/209566

[39] John Morkes, Hadyn K. Kernal, and Clifford Nass. 1999. Effects of Humor in Task-oriented Human-computer Interaction and Computer-mediated Communication: A Direct Test of SRCT Theory. *Hum-Comput Interact* 14, 4 (December 1999), 395–435. DOI:https://doi.org/10.1207/S15327051HCI1404_2

[40] Robert R. Morris, Kareem Kouddous, Rohan Kshirsagar, and Stephen M. Schueller. 2018. Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions. *J. Med. Internet Res.* 20, 6 (2018), e10148. DOI:https://doi.org/10.2196/10148

[41] A. B. C. News. Top 20 "Whitest" and "Blackest" Names. *ABC News.* Retrieved October 5, 2019 from https://abcnews.go.com/2020/top-20-whitest-blackest-names/story?id=2470131

[42] Donald A. Norman. 1994. How Might People Interact with Agents. *Commun ACM* 37, 7 (July 1994), 68–71. DOI:https://doi.org/10.1145/176789.176796

[43] M. J. O'Sullivan, P. D. Peterson, G. B. Cox, and J. Kirkeby. 1989. Ethnic populations: community mental health services ten years later. *Am. J. Community Psychol.* 17, 1 (February 1989), 17–30.

[44] SoHyun Park, Jeewon Choi, Sungwoo Lee, Changhoon Oh, Changdai Kim, Soohyun La, Joonhwan Lee, and Bongwon Suh. 2019. Designing a Chatbot for a Brief Motivational Interview on Stress Management: Qualitative Case Study. *J. Med. Internet Res.* 21, 4 (2019), e12231. DOI:https://doi.org/10.2196/12231

[45] M. Popovic, D. Milne, and P. Barrett. 2003. The scale of perceived interpersonal closeness (PICS). *Clin. Psychol. Psychother.* 10, 5 (2003), 286–301. DOI:https://doi.org/10.1002/cpp.375

[46] Byron Reeves and Clifford Ivar Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge University Press, New York, NY, US.

[47] Charles R. Ridley. 1984. Clinical treatment of the nondisclosing Black client: A therapeutic paradox. *Am. Psychol.* 39, 11 (1984), 1234–1244. DOI:https://doi.org/10.1037/0003-066X.39.11.1234

[48] Gerald L. Russell, Diane C. Fujino, Stanley Sue, Mang-King Cheung, and Lonnie R. Snowden. 1996. The Effects of Therapist-Client Ethnic Match in the Assessment of Mental Health Functioning. *J. Cross-Cult. Psychol.* 27, 5 (September 1996), 598–615. DOI:https://doi.org/10.1177/0022022196275007

[49] Carol D Ryff and Corey Lee M Keyes. The Structure of Psychological Well-Being Revisited. 9.

[50] Young June Sah and Wei Peng. 2015. Effects of visual and linguistic anthropomorphic cues on social perception, self-awareness, and information disclosure in a health website. *Comput. Hum. Behav.* 45, (April 2015), 392–401. DOI:https://doi.org/10.1016/j.chb.2014.12.055

[51] Peter Salovey and John D. Mayer. 2005. *The science of emotional intelligence.*

[52] Ari Schlesinger, Kenton P. O'Hara, and Alex S. Taylor. 2018. Let's Talk About Race: Identity, Chatbots, and AI. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–14. DOI:https://doi.org/10.1145/3173574.3173889

[53] Ryan M. Schuetzler, G. Mark Grimes, and Justin Scott Giboney. 2019. The effect of conversational agent skill on user behavior during deception. *Comput. Hum. Behav.* 97, (August 2019), 250–259. DOI:https://doi.org/10.1016/j.chb.2019.03.033

[54] Herbert W. Simons, Nancy N. Berkowitz, and R. John Moyer. 1970. Similarity, credibility, and attitude change: A review and a theory. *Psychol. Bull.* 73, 1 (1970), 1–16. DOI:https://doi.org/10.1037/h0028429

[55] Lonnie R. Snowden. 1999. African American service use for mental health problems. *J. Community Psychol.* 27, 3 (1999), 303–313. DOI:https://doi.org/10.1002/(SICI)1520-6629(199905)27:3<303::AID-JCOP5>3.0.CO;2-9

[56] Mirjam Stieger, Marcia Nißen, Dominik Rüegger, Tobias Kowatsch, Christoph Flückiger, and Mathias Allemand. 2018. PEACH, a smartphone- and conversational agent-based coaching intervention for intentional personality change: study protocol of a randomized, wait-list controlled trial. *BMC Psychol.* 6, 1 (September 2018), 43. DOI:https://doi.org/10.1186/s40359-018-0257-9

[57] S. Sue, D. C. Fujino, L. T. Hu, D. T. Takeuchi, and N. W. Zane. 1991. Community mental health services for ethnic minority groups: a test of the cultural responsiveness hypothesis. *J. Consult. Clin. Psychol.* 59, 4 (August 1991), 533–540. DOI:https://doi.org/10.1037//0022-006x.59.4.533

[58] Stanley Sue. 1977. Community mental health services to minority groups: Some optimism, some pessimism. *Am. Psychol.* 32, 8 (1977), 616–624. DOI:https://doi.org/10.1037/0003-066X.32.8.616

[59] L. K. Sussman, L. N. Robins, and F. Earls. 1987. Treatment-seeking for depression by black and white Americans. *Soc. Sci. Med. 1982* 24, 3 (1987), 187–196. DOI:https://doi.org/10.1016/0277-9536(87)90046-3

[60] Aditya Nrusimha Vaidyam, Hannah Wisniewski, John David Halamka, Matcheri S. Kashavan, and John Blake Torous. 2019. Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *Can. J. Psychiatry Rev. Can. Psychiatr.* 64, 7 (July 2019), 456–464. DOI:https://doi.org/10.1177/0706743719828977

[61] V. Venkatesh, Thong, J.Y.L., and X. Xu. 2012. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Q.* 36, 1 (2012), 157–178.

[62] Jessica Vitak. 2016. A digital path to happiness? In *The Routledge handbook of media use and well-being: International perspectives on theory and research on positive media effects*. Routledge, 274–287.

[63] Marilyn A. Walker, Janet E. Cahn, and Stephen J. Whittaker. 1997. Improvising Linguistic Style: Social and Affective Bases for Agent Personality. *ArXivcmp-Lg9702015* (February 1997). Retrieved October 5, 2019 from http://arxiv.org/abs/cmp-lg/9702015

[64] Joseph B. Walther. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Commun. Res.* 23, 1 (1996), 3–43.

[65] Joseph B. Walther and Kyle P. D'Addario. 2001. The Impacts of Emoticons on Message Interpretation in Computer-Mediated Communication. DOI:https://doi.org/10.1177/089443930101900307

[66] Henny A Westra, Adi Aviram, and Faye K Doell. 2011. Extending Motivational Interviewing to the Treatment of Major Mental Health Problems: Current Directions and Evidence. *Can. J. Psychiatry* 56, 11 (November 2011), 643–650. DOI:https://doi.org/10.1177/070674371105601102

[67] WHO. WHO | Mental disorders affect one in four people. *WHO*. Retrieved October 8, 2019 from https://www.who.int/whr/2001/media_centre/press_release/en/

[68] Embodied Conversational Agent-Based Kiosk for Automated Interviewing: Journal of Management Information Systems: Vol 28, No 1. Retrieved October 8, 2019 from https://www.tandfonline.com/doi/abs/10.2753/MIS0742-1222280102

[69] Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models, second edition by Julian J. Faraway, Chapman and Hall/CRC, Boca Raton, 2016. No. of pages: 399. Price: £63.99 (book + eBook); £44.79 (eBook). ISBN 9781498720960 - Ormerod - 2017 - Statistics in Medicine - Wiley Online Library. Retrieved October 9, 2019 from https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.7282