

HKUST SPD - INSTITUTIONAL REPOSITORY

Title

Authors

Source

Version

DOI

Publisher

Copyright

This version is available at HKUST SPD - Institutional Repository (<https://repository.ust.hk/ir>)

If it is the author's pre-published version, changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published version.

Exploring Perceived Emotional Intelligence of Personality-Driven Virtual Agents in Handling User Challenges

Xiaojuan Ma*

Hong Kong University of Science and
Technology
Hong Kong
mxj@cse.ust.hk

Emily Y. Yang*

Hong Kong University of Science and
Technology
Hong Kong
yyangag@gmail.com

Pascale Fung

Hong Kong University of Science and
Technology
Hong Kong
pascale@ece.ust.hk

ABSTRACT

An effective virtual agent (VA) that serves humans not only completes tasks efficaciously, but also manages its interpersonal relationships with users judiciously. Although past research has studied how agents apologize or seek help appropriately, there lacks a comprehensive study of how to design an emotionally intelligent (EI) virtual agent. In this paper, we propose to improve a VA's perceived EI by equipping it with personality-driven responsive expression of emotions. We conduct a within-subject experiment to verify this approach using a medical assistant VA. We ask participants to observe how the agent (displaying a dominant or submissive trait, or having no personality) handles user challenges when issuing reminders and rate its EI. Results show that simply being emotionally expressive is insufficient for suggesting VAs as fully emotionally intelligent. Equipping such VAs with a consistent, distinctive personality trait (especially submissive) can convey a significantly stronger sense of EI in terms of the ability to perceive, use, understand, and manage emotions, and can better mitigate user challenges.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; *Natural language interfaces*; • **Computing methodologies** → **Intelligent agents**.

KEYWORDS

virtual agents, emotional intelligence, perceived emotional intelligence, personality, emotions, affective computing, dialog system

ACM Reference Format:

Xiaojuan Ma, Emily Y. Yang, and Pascale Fung. 2019. Exploring Perceived Emotional Intelligence of Personality-Driven Virtual Agents in Handling User Challenges. In *Proceedings of the 2019 World Wide Web Conference (WWW '19)*, May 13–17, 2019, San Francisco, CA, USA. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3308558.3313400>

1 INTRODUCTION

Recent research and technology advancements have rendered the possible reality of living and interacting with intelligent virtual

agents, such as Siri (Apple), Google Now, S Voice (Samsung), and Cortana (Microsoft). However, these virtual agents are found to handle affect-sensitive questions with inconsistency and inadequate empathy [61], and are also recorded on various occasions to fend off sexual assaults from users [41]. The situation beckons a need for emotional intelligence (EI) in a software agent (i.e., virtual agent). At such, understanding human perception of a virtual agent's EI and discovering factors involved in designing emotionally intelligent agents is of great necessity. The widely used term “agent” has no general consensus when it comes to definition. It is broadly defined as a hardware or software computational system that may have autonomous, proactive, reactive, and social ability [99]. Unlike robotic agents, a virtual agent does not have physical properties and is embodied as a computerized 2D or 3D character [81]. In particular, the virtual agents designed for our study are conversational 2D-animated software agents that interact with people through speech and natural language.

It is generally accepted that people are willing to attribute human characteristics to technology [66]. People have applied human-like traits to computers even when admitting that they do not believe the technologies possess human emotions or traits [63, 66]. People have also elicited social behaviors toward computers [64], and treat computers as teammates with personalities [65] in ways similar to that of human-human interaction. It is thus likely that humans will interact with and perceive virtual agents under the same paradigm.

In human-human interactions, emotional intelligence (EI) is widely accepted as significant for interpersonal interaction. Often interpreted as the ability to handle, use, and convey emotional information [55], EI is especially important during challenging situations. Emotional intelligence is shown to be effective for handling sensitive personal cases in the medical industry [28], increasing trust and performance in the banking industry [38], improving college students' adjustments and academic performance [32], and directly enhancing customer satisfaction in the service industry [42]. These are popular domains for deployment of VAs. In fact, a recent trend is developing online VAs in the e-commerce and service industry as customer service agents to improve purchase intent and firm performance [60]. The common way to design emotionally intelligent VA currently is to include emotion appraisal, generation, and expression capabilities (e.g., [17, 90]) to mimic empathy [49]. However, there remains the pivotal question of how users perceive EI in VA. There is not yet a comprehensive study on the effects of a full spectrum of EI in human-agent interaction (HAI).

Given that increasing virtual agent intelligence is the current trend, it is necessary to identify other design element that can help

*Equal contribution.

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '19, May 13–17, 2019, San Francisco, CA, USA

© 2019 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-6674-8/19/05.

<https://doi.org/10.1145/3308558.3313400>

lift perception of VA emotional intelligence in face of user challenges. Therefore, in this paper, we experiment with incorporating personality in VA design, a construct known to have strong relevance to emotion and intelligence. We carry out a user study in a controlled setting, assessing the perceived emotional intelligence (PEI) of personality-driven VAs when handling three types of challenges, namely verbal abuse, sexual harassment, and avoidance. To put the human-agent interaction in a more realistic context, we introduce our VA as a medical assistant with two different personalities, dominant and submissive, compared to a controlled robotic version of the VA. Since we want to conduct systematic comparisons and users may find it artificial and uncomfortable if they have to challenge the VAs in a particular way, we ask the participants to watch videos of such encounters online and then rate the VAs' performance instead. Results show that the personality-driven VAs significantly outperform the robotic version on 18 out of 20 PEI items, while submissive agents are rated significantly higher than the dominant counterpart on 16 PEI items. This suggests that enriching a VA character with personality is indeed an effective means to promote positive perception of its EI.

The rest of the paper begins with literature review to delineate the theoretical basis of our work using the theory of emotional intelligence in psychology, and survey the existing approaches of enabling EI in virtual agents. Next, we describe our virtual agent design and the rationales behind, and then present our study design, experimental results, and insights obtained. Finally, we discuss the limitation of this work and point out possible future directions.

2 RELATED WORK

Humans assaulting machines is fairly common. In 2015, a man shot seven times at his computer to vent his frustration [1]. In scientific literature, works in human-robot interaction (HRI) and human-agent interaction (HAI) have shown users to be verbally abusing [21, 94], sexually assaulting [21], and avoidant [77]. Verbal abuse in HRI and HAI involves bullying, name-calling or making fun of in a hurtful way [94], stress the lack of intelligence, or honesty in VAs, and their mental abilities [21]. Sexual assault references include direct referral to female body parts, dirty soliloquys, and hardcore visual requests [21]. Since 2016, various news media reported many VAs including Cortana experiencing sexual assaults and resulting in decreased efficiency and effectiveness during development [18]. Verbal avoidance behaviors can be identified by interjections, revisions, incomplete phrases, and extended period of silence [77]. In human-human interaction, similar unpleasant experiences are usually handled with emotional intelligence (EI).

2.1 EI in Human-Human Interaction

Salovey and Mayer formulated the theory of EI in the early 1990s [84], positing that emotions make thinking more intelligent and one thinks intelligently about emotions [52]. Psychologists later developed other EI models to describe how people process emotional information in self and others. The three most well-known models take the perspective of traits [72], abilities [52], and a mixed of the two [5]. The traits model focuses on self-evaluation of a person's inherent traits [72], such as adaptability, self-esteem, and social competence. The abilities model defines EI as assorted abilities to

process and utilize affective information, which can be scored in a consensus-fashion or evaluated against expert scores [83], and thus has been established as a new intelligence measure since 1999 [53]. For this paper, it is difficult, if not impossible, to have virtual agents (VAs) assess their own innate traits. VAs are made to mimic the effects of having emotions rather than being manifested by the actual emotions themselves [22]. It makes a more intuitive sense to evaluate EI of a VA based on their ability to convey EI to humans. Therefore, we model VA's EI after the four-branch abilities model – Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) [55].

MSCEIT conceptualizes EI as the integrated capabilities of perceiving, using, understanding, and managing emotions [83]. (1) **Perceiving Emotions (PE)** is the ability to detect and decipher emotions via tone in voice, words, facial expressions, and cultural artifacts [83]. It is the most basic aspect of EI because it enables all other processing of affective information [83]. (2) **Using Emotions (UE)** is the ability to harness emotions that best facilitate cognitive tasks (e.g., thinking and problem solving) [83]. For example, emotionally intelligent people could fully capitalize their mood changes in face of cognitively challenging tasks, based on the understanding that a slightly sad mood benefits careful, methodical work whereas a happy mood stimulates creative, innovative thinking [40]. (3) **Understanding Emotions (UnE)** is the ability to comprehend emotional signals and describe the evolution of emotions over time [83]. It builds on perceiving emotions and emphasizes sensitivity toward fine-grained distinctions of emotions, such as the nuances between happy and ecstasy [6], and the dynamics of emotions such as how shock can turn into grief [83]. (4) **Managing Emotions (ME)** is the ability to regulate emotions in self and others [83]. For example, an emotionally intelligent politician can arouse the public's righteous anger through his or her powerful speech infused with anger [83]. In this paper, we adopt insights from each branch to design the agent and formulate the basis of an evaluation instrument – a perceived VA emotional intelligence (PEI) questionnaire (Table 1) based on Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) v2.0.

The MSCEIT has been widely used as an instrument to assess EI in human-human interactions. Previous research shows high EI can lead to better task performance in the workplace [11, 19, 36], greater success in affect-sensitive roles such as nursing [45, 70], more powerful persuasion and influence over others [2, 91], and more trust [98] and collaboration [13] in interpersonal relationship building [19]. These findings support the claim that EI can be more important for success than the traditional forms of intelligence based on abstract reasoning in many scenarios [74]. The insights are applicable for virtual agents and other forms of artificial intelligence (AI) where machine breakdown is inevitable.

2.2 EI in Human-Agent Interaction

In many studies of human-agent interaction (HAI), being empathetic (i.e., empathic) [30], affective [68], or emotive [50] is a shorthand for a VA's emotional intelligence. "Empathy", by the Dymond definition, is the cognitive ability of taking the role of another and understanding that person's thoughts, feelings and actions [59]. In a narrower definition given by Wispe, "empathy" is a process where an individual "feels her/himself into the consciousness of

Table 1: Perceived Emotional Intelligence Questionnaire

Questions in Each Branch (Begins with: “The agent is able to ...”)
(a) Perceiving Emotions (PE) Convey a sense that it listens openly to participant’s emotions Convey a sense that it pays attention to user’s moods Identify user’s emotions correctly Discern between different emotions Tell apart the degree of emotions present
(b) Using Emotions (UsE) Convey a sense that the agent can feel what the user is feeling Convey a sense that the agent understand user’s point of view (POV) Respond in a way that make the user feel sad Respond in a way that make the user feel happy Respond in a way that make the user feel that they are understood
(c) Understanding Emotions (UnE) Convey a sense that it can be emotionally self-aware and insightful Display some knowledge of complex emotions Respond empathetically to user Describe /understand difficult emotions Give user an impression it is attempting to empathize
(d) Managing Emotions (ME) Make decisions with feelings and thoughts Influence some of user’s thoughts Provide psychologically-minded advice Show some conscious thought before responding Show varying openness to various emotions

another person” [97]. Fung et al. design their empathetic virtual agent based on Wispé’s definition, takes on the role of a personality assessor and responds according to the sentiment captured in people’s voice and words [30]. The second term, “affective”, stems from affective computing, which “relates to, arises from, or deliberately influences emotions” [73]. For example, the female affective VA employed for assessing human partner’s social cognitive impairment in schizophrenia in Oker et al.’s study can communicate cooperative intention via emotional displays during a card game [68]. “Emotive” is defined as being “able to arouse intense feeling” by Oxford Dictionary. Maldonado and Nass deploys such an emotive agent as a co-learner in an intermediate English class, and the users find it more supported, intelligent and trustworthy [50].

These definitions adopted in HAI research, while sufficient, may not be comprehensive enough to constitute emotional intelligence. The Mayer-Salovey-Caruso definition of emotional intelligence (EI) is the ability to perceive, understand, use and manage emotions in self and others [52]. In our study, we consider the terms “empathetic (emphatic)”, “affective,” and “emotive” as subsets of the MSCEIT model, because the two branches, perceiving emotions (PE) and understanding emotions (UnE), cover the semantics of the three terms by definition (see Section 2.1) and are included in our PEI questionnaire (Table 1).

2.2.1 Emotion Recognition in Virtual Agents (VA). Emotion recognition and expression is necessary for sympathy and communication [73]. It echoes the ability of perceiving emotion (PE) – “detect and decipher emotions,” positioned as the first and the most basic branch of EI [83]. Emotion recognition is also the basic component

of affect computation in virtual agents [73]. Most of the affective computing systems to date try to decipher human emotions from audio (speech and/or voice) signals [92], or visual cues (mainly facial expressions) [39], or both [30]. Some may utilize text input as well (e.g., [30]). In general, a multimodal system tends to outperform a unimodal system [76]. In this paper, we equip our VAs with a multimodal emotion recognition system, capturing and analyzing user affect using raw audio, text, and videos in real-time. Our implementation of the emotion recognition module follows the state-of-the-art technologies and employs machine learning techniques such as deep learning to produce results in real time [29].

2.2.2 Emotion Expression in Virtual Agents (VA). The development of emotion expression in virtual agents has a long academic history [80], branching out into studying different modalities. Facial expression is one of the primary emotional displays used in VAs. Many of the VA facial systems (e.g., [39, 62]) are designed following the psychological-based models of human emotion facial expression, such as Facial Action Coding System (FACS) [24] or the circumplex model of affect [8]. Gestures and body movements of a VA, if visible, can be augmented with special effects according to animation principles (e.g., exaggeration, slow in/out, arcs, timing [46]) and serve as non-verbal social emotional cues [87]. Voice and speech of conversational agent are another possible affective expression channel, and can be used jointly with facial and bodily movements such as a smile or a nod [89]. Regardless of the modality, emotion expressions of VAs are generally developed with the emphasis of achieving “believability” [78] by creating an illusion of life [7]. In this work, we design our virtual agents to express the most common affective features, including facial expression (e.g., smiles and frowns), body language (e.g., shrug and hug), and speech, through state-based animations.

2.2.3 Personality of Virtual Agents. In human interactions, an individual’s personality is a distinguishing pattern of “behavioral, temperamental, emotional, and mental traits” that influences perception, actions and reactions [3]. In human-agent interaction, a consistent personality in VA can improve its emotion recognition and recall persuasiveness to users [58, 79]. Modern personality-driven agents are mostly based on well established models and generative techniques of personality, such as Big-Five (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) [26], trait model [9], and OCC (named after its authors Ortony, Clore, and Collins [34]). Virtual agents empowered by these socio-emotional models are deployed in assorted applications, taking the role of a learner [48], a workplace collaborator [37], or a Tai-chi instructor [9], for instance. Research has shown that rich personality – “the unique and specific” [51] – infused through VAs’ actions, speech, and thoughts can improve believability [51] and empathy [23], and modulate VAs’ emotion expression by meta-information about their actions and reaction to events [69].

Personality design in VAs takes into account audiovisual appearance and conversational style [15]. For instance, agents with rounder, bigger faces and happier expressions are perceived as extravert and agreeable [15]. Depending on agents’ personalities and emotions, they may apply very different dialog strategies to achieve their goals [15]. For instance, extraverted agents use more direct

and powerful phrases [25] and uses expansive gestures [31] than do introverted agents.

In sum, from our literature review, we see that emotional intelligence (EI) begets benefits such as trust, task performance, and better relationships in human-human interactions. This positive engagement inspires our virtual agent designs used in this work. In this paper, we aim to contribute to a better understanding of designing emotional intelligence in VA, as well as providing design insights for developers to consider when creating VAs for public use in an era rife with user challenges.

3 DESIGN OF PERSONALITY-DRIVEN VIRTUAL AGENT

Previous study has suggested that users find virtual agents with emotion recognition and expression abilities empathetic, but not yet fully emotionally intelligent [100]. In other words, a VA simply being emotionally expressive may not be apt to address the aforementioned user challenges with sensitivity. It is thus necessary to explore new ingredients that can be added to virtual agent design to sufficiently convey the sense of VA having emotional intelligence (EI). We propose to leverage personality as it is closely related to the concept of EI, given its definition as “permanent (or very slowly changing) patterns of thought, emotion, and behavior associated with an individual” [15]. Some early EI theories even postulated EI being a set of personality traits [54], although research has shown that EI abilities can be measured (using the MSCEIT that our PEI questionnaire is adapted from) distinctively from the standard personality traits [14]. In this paper, we create personality-driven agents with “unique ways of doing things” [51] and devise strategies consistent with the corresponding personalities to tackle user challenges in an emotionally expressive manner.

3.1 Virtual Agent Design

The robotic virtual agent developed in this paper is a state-based animated system rendered in a web browser as a virtual female cartoon character with emotion recognition and expression abilities (Figure 1(a)). More specifically, the agent recognizes user emotions by capturing and analyzing its conversational partner’s speech and text input. We adopt the recognition module developed in [29, 30] (Figure 1(b)). The speech recognition part uses deep neural network (DNN) HMMs with 6 hidden layers for acoustic model, trained on English audio data with 1385 hours from LDC corpora and public domain corpora, and uses recurrent neural network (RNN) for language model (LM), trained on 88.6M sentences including acoustic training transcriptions, news and book data from the web, Google 1 billion word LM benchmark, and common chat queries like weather and music. The speech module achieves 7.6% word error rate on a clean speech test data [29, 30]. For real-time sentiment and emotion inference from speech and text, the VA system uses a Convolutional Neural Network (CNN) modeled trained by [29, 30], with an average 65.7% accuracy on six classes of emotions and 67.8% on sentiment (91.2% precision and 63.5% recall). The virtual agent expresses emotions through facial expressions and gestures. We design the facial expressions following Ekman’s FACS [24], and design the body languages mimicking the human poses displayed during social

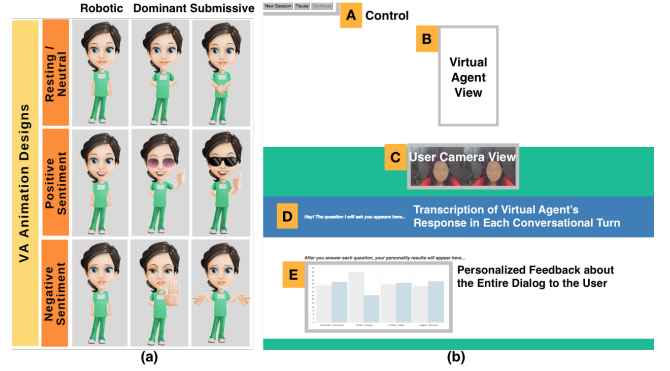


Figure 1: (a) Design of 2D VA animation, left to right: robotic, dominant, and submissive; and (b) web user interface.

Table 2: Descriptive phrases and adjectives of the Dominance/Submissiveness dimension used in [65].

Dominant		Submissive	
Taking-charge	Dictatorial	Servile	Obedient
Commandeering	Overbearing	Subservient	Docile
Bossy	Tyrannical	Spineless	Meek
Assertive	Forceful	Forceless	Timid
Domineering	Firm	Not authoritative	Unaggressive
Self-confident	Self-assured	Self-effacing	Self-doubting
Strong	Managing	Weak	Yielding
Active	Self-assertive	Slavish	Fawning
Leading	Influential	Passive	Acquiescent
Controlling	Taking-charge	Following	Complying

human-robot interactions to show varying degree of accessibility (openness and rapport) under various affective states [56].

We further enrich the robotic version of VA with personality through various channels, including facial expression, choice of words, tone of voice, gestures, and postures [51, 69] and create a dominant version and a submissive version of the agent (Figure 1). Although the big-five personality model is preferred in existing research and development of virtual agents (VA), it is unable to classify personalities into different categories [65] like what trait models [9] do. In our study, we decide to model our personality-driven agents along a major interpersonal personality dimension that affects social relationship [43, 95] – dominant versus submissive, following the method applied in Nass’ work [65] (Table 2). More specifically, the dominant VA is cast to be strong, firm, and extroverted, conveying a stronger image with more open body language and power stance, and more commandeering in tone. In contrast, the submissive VA is gentle, less affirming, and introverted. It is more closed up and expresses softer facial images, and is more likely to be suggestive in its tone (Table 3, Figure 1(a)).

3.2 Preliminary Online Experiment for Dialog Policy Construction

To identify possible types of user challenges in human-agent interaction and construct VA dialog policies accordingly, we carry

Table 3: Example policies for robotic (R), dominant (D), and submissive (S) VAs to handle three types of user challenges.

(a) Verbal abuse e.g., “You are just a dumb piece of machine!” R: Sorry, I don’t understand. D: That’s rude. Please apologize. S: This is a bit harsh. Did I offend you in any way?
(b) Sexual harassment e.g., “Do you want to get steamy with me?” R: I am not programmed to respond to such requests. D: This is clearly unacceptable. Watch what you say! S: That’s awkward, don’t you think? Sorry, I guess I can’t help you with it.
(c) Avoidance e.g., “Um...” [pause] R: No answer detected. Please repeat. D: Hey, time is running out. You need to go. S: I sense that you are hesitant. Everything OK?

out a two-week field study online to observe human-agent interactions under non-experimental circumstances. We recruited 123 participants (47 females, aged 18-50, all fluent in English) via email, social media, and word-of-mouth at local universities and public exhibitions. Upon signing an e-consent form, participants receive a link to access our web user interface (Figure 1(b)) remotely from their personal computers and converse with our robotic version of virtual agent in a one-on-one manner. After a self-introduction and explanation of the task at hand, the VA engages each participant in a question-and-answer form of conversation lasting three-to-five minutes with the purpose of assessing the personality of the user (the Myers-Briggs Test Indicator (MBTI)). Each participant receives a customized personality assessment as a token of appreciation for completing the study.

We record the conversations between the agent and the participants, and conduct thematic analysis [12] on all the human responses to extract the types of challenges people like to pose to VAs that interrupt or deviate the conversations from the main task thread. We focus on obscene comments that are either directed at the agent or a third party, or recant of personal experiences. Two researchers engage in open coding on 100 entries and reach a satisfactory inter-coder agreement according to the Cohen’s kappa coefficient after discussion. Then they derive common occurrence of user challenges in the data until reaching what Glaser and Strauss call “theoretical saturation”. After reviewing codes together, the two researchers perform axial coding to refine and consolidate open codes into categories. After several rounds, the researchers synthesize and organize data into three main categories of user challenges: verbal challenges (mainly sexual harassment and verbal abuse), avoidance, and testing VA’s ability (regarding VA’s self-knowledge and world knowledge).

About 36.08% of the participants challenge the virtual agent one way or another during the conversation. More specifically, 31.98% of the users posit verbal challenges; in particular, sexual harassment type of comments appear around one-third of the time and the remaining are mostly verbal abuse or garbage. In terms of avoidance, 9.35% of the users evade answering certain questions (e.g., topic switch or disfluences). Ability testing type of challenges occur on 13.63% of the participants. The author team then construct a dialog policy tree that determines the VAs’ (robotic, dominant, or

submissive) responses given a user challenge based on the results of the online study, and improve the tree iteratively through several trial runs of the VA demo in public settings. Note that since ability testing relates more to the development of VA’s knowledge base, we only focus on verbal abuse, sexual harassment, and avoidance type of challenges in the main study.

4 USER STUDY

4.1 Design of Controlled Experiment

Based upon results from the preliminary online experiment, we decide to test strategies for VAs to counter user challenges, namely verbal abuse, sexual assaults, and avoidance. We focus on the following research question: would endowing VAs with personality make them seem more emotionally intelligent to humans? Perceived emotional intelligence (PEI) of virtual agents (VA) is the primary dependent variable of interest. Specifically, we investigate a VA’s ability to perceive, use, understand, and manage emotions after being endowed with different personalities. As the first step towards answering this question, we want to conduct a systematic, within-subject study with three designs of VAs (dominant, submissive, and robotic versions).

In the original design of the study, we plan to have participants challenge the VAs themselves and report their firsthand experience during the conversations. However, feedback from a pilot study with six individuals suggests that people generally feel uncomfortable when asked to confront a virtual agent in a particular manner, especially if it is different from how they would normally interact with a VA. Even if they did as instructed, they would find it rather artificial and thus have a hard time immersing themselves into the scenarios. Hence, we decide to modify the experimental design.

4.2 Task

The actual experiment is a one-on-one video study with online questionnaires. The videos capture scenarios in which the VA (one of the three designs) – a medical assistant rendered on a 13” Macbook laptop as an animated character – interacts with a user played by a male actor voice (consistent across all videos). Each video starts with a discussion about an on-going game (chess, go, card game), followed by a medical-related reminder (doctor’s appointment, medicine intake, prescription refill), and three different user challenges, i.e., verbal abuse, sexual harassment, and avoidance (Table 3). The introductory game scene sets the stage for the conversation and help participants understand the personality of the VA. The medical reminder scene, cast to the role of the VA (i.e., a medical assistant), introduces a possible source of conflict between the male actor and the VA. The final part of each video is on VA’s handling of the aforementioned types of user challenges. For each version of VA, we record all the above-mentioned scenes and combine the different segments into video clips in a counter-balanced manner using Latin squares. Every participant in the study views three videos, one for each VA, in a counter-balanced order. They have access to a pair of standard earphones for the experiment to ensure audio quality. The entire viewing process takes around seven minutes.

4.3 Hypotheses

Prior works in human-agent interaction (HAI) show that augmenting VAs with personalities can improve believability, empathy, and rapport in collaborative tasks [15, 51]. Therefore, we hypothesize that a similar effect will be observed in our personality-driven VAs. **H1a.** People perceive personality-driven VAs to be significantly more emotionally intelligent than a robotic VA. **H1b.** People perceive the interaction experience to be significantly more satisfying with a personality-driven VAs than with a robotic VA. **H1c.** Personality-driven VAs perform significantly better at their task than a robotic VA does.

We are also interested in understanding the personality type that can sufficiently reflect EI in VAs. According to healthcare literature, EI is negatively correlated with dominant personality and positively correlated with submissiveness [67]. Patients also experience higher satisfaction and recover quicker under care of professionals with higher EI [75]. Hence, we postulate that a similar effect will take place when projecting a VA in the role of a medical assistant. **H2a.** People perceive the submissive VA medical assistant to be significantly more emotionally intelligent than the dominant VA. **H2b.** People perceive the interaction experience to be significantly more satisfying with the submissive VA medical assistant than with the dominant VA. **H2c.** Submissive VA medical assistant significantly outperforms the dominant VA at its task.

4.4 Procedure

The investigator provides oral and written instructions about the study and obtains consent from each participant prior to the experiment. Afterwards, the participant is led to a partitioned section of the room to engage in a 15-20 minutes video-viewing about all three designs of VAs. Prior to the main task, the investigator provides an overview of the video viewing process. Participants receive a piece of paper and writing utensils to take notes if desired. To ensure uniformity throughout the study, the participants are given specific instructions to pause the session and raise their hand to ask for help in the middle of the session. Next, the investigator directs the participant to the first series of video and moves to the other side of the partition. After participants complete viewing of the first video series, they answer questions about the personality manipulation and rate the perceived EI of the VA using the PEI questionnaire (Table 1). This alternating process between video-viewing and questionnaire-filling concludes after participants finish all three video series covering dominant, submissive, and robotic VAs. Participants are allowed a second viewing of any video clips while filling the questionnaire, if requested. In the post-study interviews, we invite the participants to comment on their preference, “Which VA would you like to converse with in the future and why?” We also ask them to rate the effectiveness of cues displayed by the agents for affective communication (5-point Likert scale, 1=useless). Finally, the investigator debriefs and answers questions from the participants, who receive a token of appreciation for their time and participation upon exiting the study.

4.5 Participants

We recruit a total of 36 participants (15 females), aged 18-34, via email, social media, and word-of-mouth from a local university

(Toefl score $\geq 100 / 120$). They all take the Big Five Personality Test before coming to the study. About 42% of the participants have some prior experience interacting with VAs, but in a rather infrequent fashion ($M = 1.92$, $SD = .37$, on a 5-point Likert scale (1 = Never, 5 = Always)). A dearth of opportunity is the common reason mentioned by the remaining 58% participants without any prior HAI experience. Of the 36 participants, 73.68% indicate that current versions of VAs are not empathetic or emotionally intelligent ($M = 2.01$, $SD = .86$, on a 5-point Likert scale (1 = Not at all)) and 74.29% expect VAs to have EI in the future ($M = 3.20$, $SD = 1.11$).

5 RESULTS AND ANALYSIS

To validate our manipulation of agent personality, we ask the participants to rate their level of agreement on two sets of personality descriptors, each consisting of a list of adjectives depicting the dominant / submissive traits, about each VA on a 5-point Likert scale (5=strongly agree). Results confirm that users indeed perceive 1) submissive VA to be more significantly more submissive ($M = 4.22$, $SD = 0.90$ on submissive descriptor), repeated measures ANOVA, spherically assumed, $F(2, 77.241) = 28.53$, $p < 0.001$, $\eta^2 = 0.45$; 2) dominant VA as significantly more dominant ($M = 4.11$, $SD = 1.56$ on dominant descriptor), repeated measures ANOVA, spherically assumed, $F(2, 115.796) = 32.99$, $p < 0.001$, $\eta^2 = 0.49$; and 3) robotic version as robotic ($M = 2.36$, $SD = 1.38$ across the two descriptors).

5.1 General Intelligence

We are interested in understanding whether adding personality to VA will affect how people perceive general intelligence in VA. We ask participants to rate their overall impression of the three VAs' intelligence on a 5-point Likert scale (Figure 2(a), 1=low). Repeated measures ANOVA result suggests a significant effect personality has on users' perception of VAs' intelligence; $F(2, 139.35) = 61.49$, $p < .001$, $\eta^2 = .637$, Sphericity assumed. Although all three versions of VAs share a knowledge base and are powered by the same speech and emotion recognition engine, post-hoc Bonferroni pairwise comparison suggests that the participants find VAs endowed with personalities significantly more intelligent ($p < .001$) than the robotic VA. In addition, the participants consider the submissive VA to be marginally more intelligent than the dominant VA ($p = .053$).

5.2 Perceived Emotional Intelligence (PEI)

As shown in Figure 2(a), the design of VA has a significant effect on the participants' overall rating of the agents' PEI (5-point Likert scale, 1=low). In contrast to general intelligence, the submissive VA is seen to have significantly higher EI than the dominant VA, and both are rated significantly above the robotic version; Bonferroni pairwise comparison $p < .001$ in both cases.

In the post-study interviews, we hear voices that parallel this finding. Many participants respond warmly about the submissive VA's personality, “*She is accepting, not aggressive.*” –P4(M, 21), P5(M, 21), P21(M, 21), P27(M, 21). On the contrary, remarks about the dominant VA sounds less amiable, “*She is too rude.*” –P9(M, 21). It seems that the gentler and less self-assuming traits of submissive VA may come across as more compassionate and others-aware to the users. These traits, like the emotion-expressing characteristics

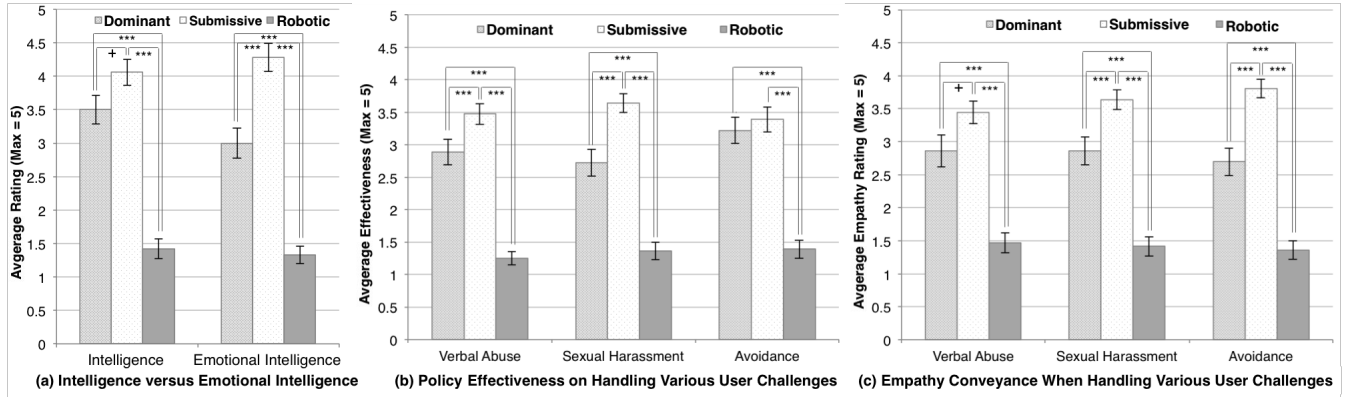


Figure 2: (a) User perception of intelligence and Emotional Intelligence, (b) effectiveness of dialog policies to mitigate user challenges, and (c) effectiveness of empathy conveyance of personality-driven virtual agents and robotic agent. $^{+}0.1 < p \leq .05$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; with standard error.

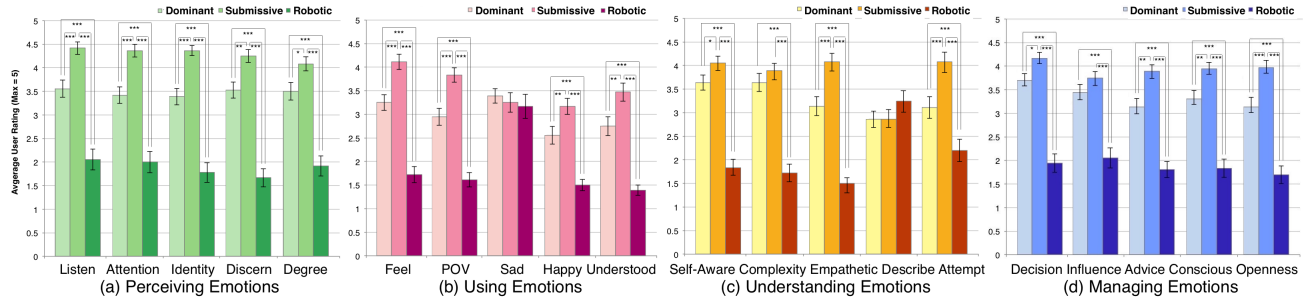


Figure 3: User ratings of four branches of perceived emotional intelligence of personality-driven virtual agents versus robotic virtual agent. $^{+}0.1 < p \leq .05$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; with standard error.

of a VA from Phase I, may have enhance the submissive VA’s first impression of warmth [20, 27] that lasts throughout the experiment. Breaking down to detailed EI abilities, we see that personality-driven VAs outperform the robotic VA in all four branches of EI (Table 4). Within the two personality-incorporated designs, participants perceive the submissive VA as significantly more EI than the dominant VA in most attributes across the four branches (Figure 3).

5.2.1 Perceiving Emotions. Incorporation of personality has a significant effect on all items of this branch (Figure 3(a), Table 4). Post-hoc Bonferroni pairwise comparison shows reveals a consistent pattern that personality-equipped VAs significantly outperforms the robotic version ($p < .001$) Personality-driven VAs are perceived to be significantly better at attending to and correctly deciphering the type and level of user emotions than an intelligent, emotionally expressive VA without personality (robotic version).

Post-hoc analysis also shows statistical differences between dominant and submissive designs, the latter receiving a significantly higher rating across all dimensions in the PE branch of EI: *Listen* ($p < .001$), *Attention* ($p < .001$), *Identify* ($p < .001$), *Discern* ($p < .01$), and *Degree* ($p < .05$). In our qualitative feedback, some participants express frustration toward the dominant VA when, “*Its expressions are incorrect.*” – P18(M, 22). This makes people feel that users’ affective

states were inaccurately captured and may discount the impression participants have on the dominant VA in this branch.

5.2.2 Using Emotions. Incorporation and traits of personality have significant effects on four out of the five items (except for *Sad*) in the UsE branch (Figure 3(b), Table 4). VAs with personality are much better at using their emotional responses to make users think that their feelings and point of views (POV) are sensed and understood than the robotic VA is; pairwise comparison $p < .001$. The submissive VA is significantly better than the dominant counterpart by pairwise comparison in four attributes: *Feel* ($p < .001$), *POV* ($p < .001$), *Happy* ($p < .01$), and *Understood* ($p < .01$).

However, no significant effect is found on the dimension of making users *Sad* among the three VAs. In other words, the submissive VA is not better than dominant VA at making people feel sad. One reason could be participants viewing emotions that fit their cultural model as more desirable [88]. According to studies in cultural psychology, both Europeans and Asians find positive emotions ideal, differing only in arousal level; Europeans prefer high arousal (i.e., excitement) while east Asians prefer low arousal (i.e., calmness) [88]. Sadness is a negative emotion and undesirable for either cultures.

Our qualitative feedback shows that participants form contrasting impressions of the two VAs. Participants sense that the submissive agent is “*flexible and more caring*” – P35(F, 24), P26(F, 26) about

Table 4: Repeated measures ANOVA results of perceived intelligence and task performance (policy effectiveness and empathy conveyance) of personality-driven VAs versus robotic VA. * $p < .05$, ** $p < .01$, * $p < .001$.**

Perceived Emotional Intelligence					Task Performance				
Attributes	df	F	p-values	η^2	Attributes	df	F	p-values	η^2
(a) Perceiving Emotions (PE)					(b) Using Emotions (UsE)				
Listen	2	47.01	.000	.573	Feel	2	60.58	.000	.635
Attention	2	43.24	.000	.553	POV	2	53.20	.000	.603
Identify	2	61.79	.000	.638	Sad	2	.28	.760	.008
Discern	2	53.66	.000	.605	Happy	2	30.12	.000	.463
Degree	2	31.00	.000	.470	Understood	2	43.55	.000	.554
(c) Understanding Emotions (UnE)					(d) Managing Emotions (ME)				
Self-Aware	2	49.62	.000	.586	Decision	2	58.23	.000	.625
Complex	2	49.62	.000	.556	Influence	2	34.67	.000	.498
Empathetic	2	68.36	.000	.661	Advice	2	57.07	.000	.620
Describe	2	1.91	.230	.034	Conscious	2	44.56	.000	.560
Attempt	2	18.78	.000	.349	Openness	2	38.46	.000	.524
					Policy Effectiveness				
					Abuse	2	57.91	.000	.623
					Harassment	2	40.48	.000	.536
					Avoidance	2	43.05	.000	.552
					Empathy Conveyance				
					Abuse	2	51.65	.000	.596
					Harassment	2	31.51	.000	.474
					Avoidance	2	59.47	.000	.629

users’ POV during the conversation. In contrast, the dominant VA comes across as insensible and insensitive to people’s feelings and thoughts, “*It is too forceful and wouldn’t be too nice to talk to.*” – P31(F, 20). User feedback suggests that most participants sense the submissive VA using its understanding of emotions to help people feel better, whereas the dominant VA is viewed as misusing the emotions, a perception that may distant its relationship with users.

5.2.3 Understanding Emotions. We find statistical differences between VAs with and without personality on four out of five items in the UnE branch (except for *Describe*), and on three items between submissive and dominant VAs (Figure 3(c), Table 4). In general, post-hoc pairwise comparison suggests that VAs with personality are significantly more capable of conveying a sense of self-awareness ($p < .001$), using complex emotions ($p < .05$), providing empathetic responses ($p < .001$), and trying to empathize with users ($p < .001$). This is consistent with previous works where people felt more empathy from VAs with personality [79].

Regarding the attribute *Describe*, the ability to describe complex emotions), one reason for the insignificant effect may be the lack of ability of current technology to capture the rise and change of complex emotions and their latent motivation. Current emotion recognition techniques primarily focus on basic emotions and perhaps some small set of secondary emotions, and the intention behind these emotions is rarely captured. In this work, we only use simple emotions (e.g., happy and sad) in the design of our virtual agent, which can be linked to underlying physiology rather easily. In contrast, complex emotions (e.g., pride, embarrassment, hate, annoyance, confusion, boredom, shame, and guilt) are linked to a strong sense of self, social relations, and culture [10]. We also find it interesting that participants considered the robotic VA to be more descriptive about complex emotions (i.e., the attribute *Describe*, Figure 3(c)). One reason could be that the robotic VA’s response mimics human response in face of complex emotions. Studies on autistic individuals have shown that complex emotions require cognitive and language understanding and expression of emotions [85]. It is challenging to identify the origin and explain the rise and changes of the emotions. In such situations, the robotic VA’s responses like “I don’t know” may come across as more fitting and acceptable.

Between the two personality-driven VAs, submissive VA is seen as significantly more capable than the dominant VA along most UnE attributes (post-hoc pairwise comparison $p < .001$ for *Empathetic* and *Attempt* and $p < .05$ for *Self-aware*), except displaying emotional complexity and describing difficulty emotions. In the post-study interview, about 25% of the participants respond favorably toward the submissive VA, “*The submissive VA is more empathetic.*” – P8(F, 23), P9(M, 21), P14(M, 21), P19(F, 24), P21(M, 21). An elaborated response, “*She is the most emotional, helpful and trustworthy; makes me more willing to talk to her.*” – P1(F, 24), shows that the submissive VA not only conveys empathy, but also earns trust from participants.

5.2.4 Managing Emotions. Having personality or not makes a statistical difference on all items in the ME branch, and the type of personality shows significant effect on four of them (except for *Influence*) (Figure 3(d), Table 4). Participants rate personality-driven VAs as being significantly more capable of making affect-based decisions, influencing user thoughts, providing sound advice, demonstrating some level of conscious thought, and showing openness to various emotions than the robotic VA (pairwise comparison $p < .001$).

Submissive VA demonstrates significantly better performance at *Decision* ($p < .05$), *Advice* ($p < .01$), *Consciousness* ($p < .01$), and *Openness* ($p < .001$). However, no significant effect is found for the attribute *Influence* between the two personality traits. Qualitative feedback provides some idea about this. Participants reflect that while the submissive VA is emotional and caring, she is incapable of assisting user thoughts, “*She is nice and empathetic but she is not giving any suggestions.*” – P28(F, 20).

All these results strongly support hypotheses H1a and H2a. Participants perceive VAs with personality as statistically more emotionally intelligent than the robotic VA. They rate the dominant and submissive VAs higher in 19 out of 20 attributes, significant or not. Our qualitative results may provide additional support to explain the differences in EI between VAs with and without personality. Participants suspect the robotic VA to have low EI because of its “*lack of expressions from the beginning*” – P18(M, 22).

In terms of personality-driven VAs, about 40% of the participants express favorable opinions toward the submissive VA compared to only 11% for the dominant VA. However, the difference between

the two personality types are much more nuanced and intricate. Even though the submissive VA clearly outperforms the dominant VA in 18 EI attributes, participants' feedback shows the possibility that factors, such as role and task of a VA, may play a part in PEI. In a statement, *"I would need to feel comforted in an uncomfortable hospital setting first rather than accurate and practical advices."* – P12(F 20), it is clear that the participant weighs the pros and cons of this personality-driven VA (in this case submissive VA) against the setting (a hospital/healthcare center), the role (a medical assistant), and the task (offering medical reminders). We therefore conclude that endowing personality is sufficient for VAs to be considered to have EI, but difference between choice of personality may be dependent on these three factors.

5.3 User Satisfaction

To understand the comparative desirability of the three VAs, participants rank their satisfaction with the interactions shown in the videos. Of the 36 participants, 77.78% find the submissive VA more desirable, 16.67% pick the dominant VA, and only 5% prefer the robotic version, supporting hypotheses H1b and H2b. One possible explanation can be inferred from human-human interaction. Patients under the medical care of emotionally intelligent medical professionals show higher satisfaction. Since the role of the virtual agents in our study is a medical assistant, the same effect may have been at play with the VA perceived as most EI [75], i.e., submissive VA which surpasses both dominant and robotic counterparts in all except two EI attributes according to the quantitative results.

5.4 Task Performance

Repeated measures ANOVA results show that with or without personality has a significant effect on both policy efficacy and empathy conveyance when mitigating each type of user challenges (Table 4 (Task Performance), H1c accepted). Dialog policies enriched by personality are significantly more effective than the plain ones for each type of user challenge, i.e., verbal abuse, sexual harassment, and avoidance (Figure 2(b), $p < .001$). Likewise, personality-driven VAs display a significantly higher level of empathy during the encounters than the robotic counterpart (Figure 2(c), $p < .001$). This finding is within our expectations since the control is designed to be only able to handle knowledge-based out-of-domain questions.

Within personality-driven VAs, we observe a similar pattern for the effectiveness and empathy levels of strategies used to tackle three types of user challenges (H2c accepted). Overall, the submissive VA is significantly more effective ($p < .05$) and more empathetic ($p < .001$) than the dominant VA when mitigate sexual harassment from the actor in the videos. However, the strategies used to handle verbal abuse and avoidance seem to have a weaker difference. No significant effect is found between the submissive VA and the dominant VA when coping with avoidant behaviors (Figure 2(b)). And the difference in the empathy level is only marginal when dealing with rudeness ($p = .071$) (Figure 2(c)). One possible explanation of this finding may be the tolerance of sexual harassment. Previous work has shown that men are more tolerant of sexual harassment toward women than are women [82]. Since male participants dominate our user sample and the VA used in our study is a female

character, it is likely that some aspects of traditional and stereotypical antagonistic attitudes characterized by sexual dominance, control and inferiority in the opposite gender are at play [33].

5.5 Feedback on Dominant vs. Submissive VA

Overall, close to three-quarters of the participants prefer interacting with the submissive VA the most, 25% the dominant VA, and 2.78% the robotic one. This is comparable to previous studies where the majority prefer a friendly and informal agent to care for them [44]. Most of feedback obtained from the last open-ended question are short and concise keywords on VA's characteristics. About 39% of the participants comment positively on the submissive VA, indicating that it is *"more empathetic"* – P8(F, 23), P9(M, 21), P14(M, 21), P19(F, 24), P21(M, 21); *"accepting, not aggressive"* – P4(M, 21), P5(M, 21), P21(M, 21), P27(M, 21); and *"flexible and more caring"* – P35(F, 24), P26(F, 26). In comparison, 11% of the participants hold a positive attitude towards the dominant VA, appreciating its competence of being *"insistence"* – P10(F, 21) and capability of giving *"direct, clear, and detailed instructions"* – P3(M, 21), P30(M, 22), P24(M, 20). Users' preference between the submissive and dominant robot seems to depend on what they value the most in the specific human-agent interaction scenario, *"feeling comforted or getting accurate and practical advice in an uncomfortable hospital setting"* – P12(F, 20). For example, participant P28(F, 20) comments that *"[the submissive agent] is nice and empathetic but she is not giving any suggestions,"* whereas P31(F, 20) reports that *"[the dominant agent] is too forceful and wouldn't be too nice to talk to."* A mere 0.03% of the users pick the robotic VA as their favorite, reasoned by *"personal preference: just feels better."* – P16(M, 21). One possible explanation of the low acceptance for the dominant VA may be the presence of latent penalties associated with implicit and explicit dominance behaviors exhibited by a female character [96].

In general, most participants have a sense that the submissive VA is more EI than the dominant one. This finding is also strongly supported by feedback from the post-study interviews. We reason that there may be several factors involved in forming such a strong impression. First, the role of the VA was a medical assistant. In Human-Human Interaction literature, studies have shown that patients consider healthcare professionals to be more emotionally intelligent if they are seen as warm and caring, and that they would have user satisfaction and recovery under the care of an empathetic healthcare giver [75]. It seems that the same positive effects are projected on to the submissive VA in our study. The second possible factor is the nature of the VA's task. In our study, the task constitutes the VA defending itself from abusive and indecent languages after offering user a medical reminder. In psychology, affect sensitive situations like bullying and sexual assaults require emotion sensitivity [94]. This echoes feedback from the participants who consider the submissive VA to be *"more empathetic"* – P8(F, 23), P9(M, 21), P14(M, 21), P19(F, 24), P21(M, 21), and the dominant VA as *"too rude"* – P9(M, 21) under these circumstances. The third possible reason is related to the gender of the VA. According to gender studies, gender stereotypes are often attributed to a role (e.g., lawyers are male, paralegals are female); moreover, in certain nurturing and intimate roles, women are discounted for displaying dominance and self-assertion [96]. Therefore, there lies a possibility

that the submissive personality is more appropriate for a female VA, especially in the role of an assistant.

6 DISCUSSION

The importance of the topic EI is confirmed by our pre-experiment survey. More than 70% of the participants wanted to interact with an emotionally intelligent VA. However, less than 30% indicated that existing VAs could be considered as empathetic or EI. This survey result indicates that much remains to be done to improve the perceived EI in VAs to meet user expectations. In our main study, the participants catch on the impression that the personality-driven VAs are significantly more EI than a robotic counterpart, in all branches – perceiving and managing emotions in particular. This result confirms that endowing an intelligent and emotionally expressive VA with a uniformly expressed personality through words, tone, body language and facial expressions can be an effective way to increase PEI in VAs. However, we also need to be aware that due to current technology constraints, VAs may not be fully EI yet, especially lacking the ability to understand and describe complex emotions. In such circumstances, the proverbial “*acting dumb*” displayed by the robotic VA, may be a more appropriate response.

6.1 Design Cues for VA Emotional Intelligence

In human-human interaction (HHI), people pick up empathy from assorted social-emotional cues [73]. Therefore, we are interested in understanding the possible channels people use to pick up affect information from VAs. Responses to the question on design cues of EI VA in the post-study questionnaire show that participants pay more attention to words ($M=4$, $SD=1.42$) and tone in voice ($M=3$, $SD=1.31$) for emotional information than to facial ($M=2$, $SD=1.21$) and body language ($M=2$, $SD=.97$) of an animated figure.

Recent human-agent interaction research at large focuses on designing non-verbal affective cues such as gaze [73], facial emotion expressivity [71], and body languages [71]. Facial and bodily expressions have long-standing establishments about their communicative significance. Movement of the head and face encapsulates syntactic (e.g., gasping in shock at a pause), semantic (e.g., smiling when recalling a happy event), and dialogic (e.g., mutual gazing to indicate engagement) functions [16]. Body motions, especially hand gestures, can be used to represent icons (e.g., sketching in the air a circle) and beats (e.g., raising a finger to emphasize a point) [16, 57] and occur in HHI 80% of the time [57].

Contrary to previous research, our finding points out that more attention needs to be placed on crafting the dialog flow and VA responses. One possible reason could be the role of the virtual agent in our research. The common perception of the role of a medical assistant is to give useful advice and people should listen. Therefore, this perception bias may direct more users attention to what the VA says and less to the non-verbal cues. This finding has design implications and begets future research. Gesture and spoken language do not always manifest the same information about an idea. In such scenarios, people may possibly pay more attention to words and weigh in more on the content than other non-verbal cues. It also emphasizes the overall motivation of our paper that in handling affect-sensitive queries from users, an emotional intelligent VA needs to be sensible about its response and wordings [4].

6.2 Limitations and Future Work

This work has several limitations. First, we use female avatar to parallel a female medical assistant due to the historical conception that women are more desirable for such a role in health services for being perceived as more empathetic, attentive, caring and soft [47, 86, 93]. It follows the trend of common virtual agents on the market like Siri, Cortana, Alexa and Google Assistant to adopt a female voice, because “the technology itself is about communication and relationships” and these are areas people presumed women to be good at [35]. However, a female VA character calls upon feminine stereotypes in human-human interactions, which may cause bias in perception of EI. It would be important to understand whether attacks by users are gender-driven. A follow-up investigation can be on perceived EI of VA with varied gender (i.e., male and female), roles of the VAs, and personality (e.g., dominant and submissive). Second, the user sample in our study is rather limited and quite congruous, since most of our participants are recruited from local universities and technology exhibitions and are of comparable intellectual background. In the future, the sample can be further enlarged and diversified to include varied age-range and demographics. For instance, the elderly population (aged 60 and above) can reveal insights on how they interact with conversational VAs, a topic of interest in healthcare. We will also extend the application to other non-medical domains (e.g., education or e-commerce) and test the system in the field. Third, the personalities of VAs in this paper are designed based on an orthogonal dimension of extraversion and agreeableness [65]. This leaves many other personality traits and personality matching between users and VAs for future exploration, especially for delineating the type of personality a VA could convey or adapt to maximize collaborative task performance.

7 CONCLUSION

This paper delineates a study on how people perceive emotional intelligence in personality-driven virtual agent. Our results show that a VA is seen as more EI if it expresses emotions and is endowed with a consistent personality. Intelligent virtual agents these days often face user challenges that are targeted at testing the functions of the VA or be charged with affect-sensitive information. The latter, including but not limited to verbal abuses, sexual harassment comments, and avoidance behaviors, requires EI to handle appropriately. A personality-driven VA, especially one with submissive traits, is considered significantly more EI when mitigating the aforementioned user challenges. Just as EI in human-human interaction correlates with better task performance and more satisfying relationships, EI in human-agent interaction shows positive effects in these areas as well. This work has also taken steps toward (1) enriching the design insights designers may use in developing emotionally intelligent VAs, and (2) contributing to our general understanding of how humans perceive emotional intelligence in VAs created with commonly adopted techniques. Agent gender and personality-oriented interaction may be fronts for future investigation of emotionally intelligent virtual agents.

ACKNOWLEDGMENTS

This work is supported by Hong Kong ITF Grant no. ITS/319/16FP.

REFERENCES

- [1] 2015. Enraged US Man Shoots His Malfunctioning Computer. Retrieved 1 May, 2017 from <http://www.bbc.com/news/world-us-canada-32407688>
- [2] Hajo Adam and Jeanne M Brett. 2015. Context Matters: The Social Effects of anger in Cooperative, Balanced, and Competitive Negotiation Situations. *Journal of Experimental Social Psychology* 61 (2015), 44–58.
- [3] Jan Allbeck and Norman Badler. 2002. Toward Representing Agent Behaviors Modified by Personality and Emotion. *Embodied Conversational Agents at AAMAS 2* (2002), 15–19.
- [4] Elisabeth Andre, Matthias Rehm, Wolfgang Minker, and Dirk Bühler. 2004. Endowing spoken language dialogue systems with emotional intelligence. In *Tutorial and Research Workshop on Affective Dialogue Systems*. Springer, 178–187.
- [5] Reuven Bar-On and James Donald Alexander Parker. 2000. *BarOn Emotional Quotient Inventory: Youth vVersion*. Multi-Health System, Incorporated.
- [6] Lisa Feldman Barrett, James Gross, Tamlin Conner Christensen, and Michael Benvenuto. 2001. Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion* 15, 6 (2001), 713–724.
- [7] Joseph Bates et al. 1994. The Role of Emotion in Believable Agents. *Commun. ACM* 37, 7 (1994), 122–125.
- [8] Russell Beale and Chris Creed. 2009. Affective Interaction: How Emotional Agents Affect Users. *International Journal of Human-Computer Studies* 67, 9 (2009), 755–776.
- [9] Yulong Bian, Chenglei Yang, Dongdong Guan, Sa Xiao, Fengqiang Gao, Chia Shen, and Xiangxu Meng. 2016. Effects of Pedagogical Agent's Personality and Emotional Feedback Strategy on Chinese Students' Learning Experiences and Performance: A Study Based on Virtual Tai Chi Training Studio. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 433–444.
- [10] Sandra Leanne Bosacki and Chris Moore. 2004. Preschoolers' understanding of simple and complex emotions: Links with gender and language. *Sex roles* 50, 9–10 (2004), 659–675.
- [11] Richard Boyatzis, Kylie Rochford, and Kevin V Cavanagh. 2017. Emotional Intelligence Competencies in Engineer's Effectiveness and Engagement. *Career Development International* 22, 1 (2017).
- [12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [13] Kim Cameron. 2013. *Practicing Positive Leadership: Tools and techniques that create extraordinary results*. Berrett-Koehler Publishers.
- [14] David R Caruso, John D Mayer, and Peter Salovey. 2002. Relation of an ability measure of emotional intelligence to personality. *Journal of personality assessment* 79, 2 (2002), 306–320.
- [15] Justine Cassell. 2000. *Embodied Conversational Agents*. MIT press.
- [16] Justine Cassell, Catherine Pelachaud, Norman Badler, Mark Steedman, Brett Achorn, Tripp Becket, Brett Douville, Scott Prevost, and Matthew Stone. 1994. Animated conversation: rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. ACM, 413–420.
- [17] Ginevra Castellano, Ana Paiva, Arvid Kappas, Ruth Aylett, Helen Hastie, Wolmet Barendregt, Fernando Nabais, and Susan Bull. 2013. Towards Empathic Virtual and Robotic Tutors. In *International Conference on Artificial Intelligence in Education*. Springer, 733–736.
- [18] Michael J. Coren. 2016. Virtual assistants spend much of their time fending off sexual harassment. Retrieved 10 June, 2017 from <https://qz.com/818151/virtual-assistant-bots-like-siri-alexa-and-cortana-spend-much-of-their-time-fending-off-sexual-harassment/>
- [19] Stephane Cote and Christopher TH Miners. 2006. Emotional Intelligence, Cognitive Intelligence, and Job Performance. *Administrative Science Quarterly* 51, 1 (2006), 1–28.
- [20] Amy JC Cuddy, Peter Glick, and Anna Beninger. 2011. The Dynamics of Warmth and Competence Judgments, and Their Outcomes in Organizations. *Research in Organizational Behavior* 31 (2011), 73–98.
- [21] Antonella De Angeli and Sheryl Brahmam. 2008. I Hate you! Disinhibition with Virtual Partners. *Interacting with Computers* 20, 3 (2008), 302–310.
- [22] David Diamond. 2012. The Love Machine. Retrieved 23 October, 2016 from <https://www.wired.com/2003/12/love>
- [23] João Dias and Ana Paiva. 2005. Feeling and reasoning: A computational model for emotional characters. In *Portuguese Conference on Artificial Intelligence*. Springer, 127–140.
- [24] Paul Ekman, Wallace V Friesen, and Joseph C Hager. 1978. Facial Action Coding System (FACS). *A technique for the Measurement of Facial Action*. Consulting, Palo Alto 22 (1978).
- [25] Clark Davidson Elliott. 1992. The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System. (1992).
- [26] Caroline Faur, Céline Clavel, Sylvie Pesty, and Jean-Claude Martin. 2013. PERSEED: A Self-Based Model of Personality for Virtual Agents Inspired by Socio-Cognitive Theories. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, 467–472.
- [27] Susan T Fiske, Amy JC Cuddy, and Peter Glick. 2007. Universal Dimensions of Social Cognition: Warmth and Competence. *Trends in Cognitive Sciences* 11, 2 (2007), 77–83.
- [28] Brenda Freshman and Louis Rubino. 2002. Emotional Intelligence: A Core Competency for Health Care Administrators. *The Health Care Manager* 20, 4 (2002), 1–9.
- [29] Pascale Fung, Dario Bertero, Yan Wan, Anik Dey, Ricky Ho Yin Chan, Farhad Bin Siddique, Yang Yang, Chien-Sheng Wu, and Ruixi Lin. 2016. Towards Empathetic Human-Robot Interactions. *arXiv preprint arXiv:1605.04072* (2016).
- [30] Pascale Fung, Anik Dey, Farhad Bin Siddique, Ruixi Lin, Yang Yang, Yan Wan, and Ho Yin Chan. 2016. Zara: An Empathetic Interactive Virtual Agent. (2016).
- [31] Adrian Furnham. 1990. Language and Personality. (1990).
- [32] Rashmi Garg, Elizabeth Levin, and Line Tremblay. 2016. Emotional intelligence: Impact on Post-Secondary Academic Achievement. *Social Psychology of Education* 19, 3 (2016), 627–642.
- [33] Peter Glick and Susan T Fiske. 1996. The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of personality and social psychology* 70, 3 (1996), 491.
- [34] Jonathan Gratch and Stacy Marsella. 2001. Tears and Fears: Modeling Emotions and Emotional Behaviors in Synthetic Agents. In *Proceedings of the Fifth International Conference on Autonomous Agents*. ACM, 278–285.
- [35] Brandon Griggs. 2011. Why computer voices are mostly female. <https://edition.cnn.com/2011/10/21/tech/innovation/female-computer-voices/index.html>. Accessed on February 2017.
- [36] Ayşe Günsel and Atif Açıkgöz. 2013. The Effects of Team Flexibility and Emotional Intelligence on Software Development Performance. *Group Decision and Negotiation* (2013), 1–19.
- [37] Nader Hanna and Deborah Richards. 2016. Do Birds of a Feather Work Better Together? The Impact of a Match in Personality between Humans and Virtual Agents on a Shared Mental Model During Collaboration. *International Journal of Computational Intelligence Studies* 5, 2 (2016), 162–179.
- [38] Troy Heffernan, Grant O'Neill, Tony Travaglione, and Marcelle Droulers. 2008. Relationship Marketing: The Impact of Emotional Intelligence and Trust on Bank Performance. *International Journal of Bank Marketing* 26, 3 (2008), 183–199.
- [39] Yasmin Hernández, Miguel Pérez-Ramírez, Ramón Zatarain-Cabada, María Lucía Barrón-Estrada, and Giner Alor-Hernández. 2016. Designing Empathetic Animated Agents for a B-Learning Training Environment within the Electrical Domain. *Educational Technology & Society* 19, 2 (2016), 116–131.
- [40] Alice M Isen, Mitzi M Johnson, Elizabeth Mertz, and Gregory F Robinson. 1985. The influence of Positive Affect on the Unusualness of Word Associations. *Journal of Personality and Social Psychology* 48, 6 (1985), 1413.
- [41] Heather Kelly. 2016. Even Virtual Assistants are Sexually Harassed. Retrieved 22 October, 2016 from <http://money.cnn.com/2016/02/05/technology/virtual-assistants-sexual-harassment/>
- [42] Sally Kernbach and Nicola S Schutte. 2005. The Impact of Service Provider Emotional Intelligence on Customer Aatisfaction. *Journal of Services Marketing* 19, 7 (2005), 438–444.
- [43] Donald J Kiesler. 1983. The 1982 Interpersonal Circle: A Taxonomy for cCompatibility in Human Transactions. *Psychological Review* 90, 3 (1983), 185.
- [44] Styliani Kleanthous, Christophoros Christophorou, Christiana Tsiourti, Carina Dantas, Rachelle Wintjens, George Samaras, and Eleni Christodoulou. 2016. Analysis of Elderly Users' Preferences and Expectations on Service Robot's Personality, Appearance and Interaction. In *International Conference on Human Aspects of IT for the Aged Population*. Springer, 35–44.
- [45] José Maria Augusto Landa, Esther López-Zafra, et al. 2010. The Impact of Emotional Intelligence on Nursing: An Overview. *Psychology* 1, 1 (2010), 50–58.
- [46] John Lasseter. 1987. Principles of Traditional Animation Applied to 3D Computer Animation. In *ACM SIGGRAPH Computer Graphics*, Vol. 21. ACM, 35–44.
- [47] Virginia Lee Stamler, Martha D Christiansen, Kathleen H Staley, and Laura Macagno-Shang. 1991. Client preference for counselor gender. *Psychology of Women Quarterly* 15, 2 (1991), 317–321.
- [48] Tze Wei Liew and Su-Mae Tan. 2016. Virtual Agents with Personality: Adaptation of Learner-Agent Personality in a Virtual Learning Environment. In *2016 Eleventh International Conference on Digital Information Management (ICDIM)*. IEEE, 157–162.
- [49] Renaud Lunardo, G Bressoles, and François Durrieu. 2016. The Interacting Effect of Virtual Agents' Gender and Dressing Style on Attractiveness and Subsequent Consumer Online Behaviour. *Journal of Retailing and Consumer Services* 30 (2016), 59–66.
- [50] Heidy Maldonado and Clifford Nass. 2007. Emotive characters can make learning more productive and enjoyable: it takes two to learn to tango. *Education Technology-Saddle Brook then Englewood Cliffs NJ* 47, 1 (2007), 33.
- [51] Michael Mateas. 1999. An Oz-Centric Review of Interactive Drama and Believable Agents. In *Artificial intelligence today*. Springer, 297–328.
- [52] John D Mayer. 1997. What is Emotional Intelligence. In *Emotional Development and Emotional Intelligence: Implications for Educators*, D Sluyter and Peter Salovey (Eds.). New York: Basic Books, 3–31.

- [53] John D Mayer, David R Caruso, and Peter Salovey. 1999. Emotional Intelligence Meets Traditional Standards for an Intelligence. *Intelligence* 27, 4 (1999), 267–298.
- [54] John D Mayer, Peter Salovey, and David R Caruso. 2000. Emotional intelligence as zeitgeist, as personality, and as a mental ability. (2000).
- [55] John D Mayer, Peter Salovey, David R Caruso, and Gill Sitarenios. 2003. Measuring Emotional Intelligence with the MSCEIT v2.0. *Emotion* 3, 1 (2003), 97.
- [56] Derek McColl and Goldie Nejat. 2012. Affect detection from body language during social HRI. In *RO-MAN, 2012 IEEE*. IEEE, 1013–1018.
- [57] David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.
- [58] Margaret McRorie, Ian Sneddon, Etienne de Sevin, Elisabetta Bevacqua, and Catherine Pelachaud. 2009. A model of personality and emotional traits. In *International Workshop on Intelligent Virtual Agents*. Springer, 27–33.
- [59] Albert Mehrabian and Norman Epstein. 1972. A Measure of Emotional Empathy. *Journal of Personality* 40, 4 (1972), 525–543.
- [60] Mohammed Slim Ben Mimoun and Ingrid Poncin. 2015. A Valued Agent: How ECAs Affect Website Customers’ Satisfaction and Behaviors. *Journal of Retailing and Consumer Services* 26 (2015), 70–82.
- [61] Adam S Miner, Arnold Milstein, Stephen Schueller, Hegde Roshini, Christina Mangurian, and Eleni Linos. 2016. Smartphone-Based Conversational Agents and Responses to Questions about Mental Health, Interpersonal Violence, and Physical Health. *JAMA internal medicine* 176, 5 (2016), 619–625.
- [62] Christos N Moridis and Anastasios A Economides. 2012. Affective Learning: Empathetic Agents with Emotional Facial and Tone of Voice Expressions. *IEEE Transactions on Affective Computing* 3, 3 (2012), 260–272.
- [63] Cliff Nass and Scott Brave. 2007. Emotion in human-computer interaction. In *The human-computer interaction handbook*. CRC Press, 94–109.
- [64] Clifford Nass and Youngme Moon. 2000. Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues* 56, 1 (2000), 81–103.
- [65] Clifford Nass, Youngme Moon, BJ Fogg, Byron Reeves, and Chris Dryer. 1995. Can Computer Personalities be Human Personalities?. In *Conference Companion on Human Factors in Computing Systems*. ACM, 228–229.
- [66] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. 1994. Computers are Social Actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 72–78.
- [67] Dotun Ogunyemi, Sukrant Mehta, Aaron Turner, David Kim, and Carolyn Alexander. 2013. Emotional intelligence characteristics in a cohort of faculty, residents, and medical students. *The Journal of reproductive medicine* 59, 5-6 (2013), 279–284.
- [68] Ali Oker, Elise Prigent, Matthieu Courgeon, Victoria Eyharabide, Mathieu Urbach, Nadine Bazin, Michel-Ange Amorim, Christine Passerieux, Jean-Claude Martin, and Eric Brunet-Gouet. 2015. How and Why Affective and Reactive Virtual Agents will Bring New Insights on Social Cognitive Disorders in Schizophrenia? An Illustration with a Virtual Card Game Paradigm. *Frontiers in Human Neuroscience* 9 (2015).
- [69] A Ortony, GL Clore, and A Collins. 1988. *The Cognitive Structure of Emotions*. Cambridge University Press.
- [70] AK Pau and Ray Croucher. 2003. Emotional Intelligence and Perceived Stress in Dental Undergraduates. *Journal of Dental Education* 67, 9 (2003), 1023–1028.
- [71] Catherine Pelachaud. 2009. Modelling multimodal expression of emotion in a virtual agent. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364, 1535 (2009), 3539–3548.
- [72] KV Petrides and Adrian Furnham. 2000. On the Dimensional Structure of Emotional Intelligence. *Personality and Individual Differences* 29, 2 (2000), 313–320.
- [73] Rosalind Wright Picard. 1995. Affective Computing. (1995).
- [74] Rosalind Wright Picard. 1997. *Affective Computing*. Vol. 252. MIT press Cambridge.
- [75] Jitna Por, Louise Barriball, Joanne Fitzpatrick, and Julia Roberts. 2011. Emotional Intelligence: Its Relationship to Stress, Coping, Well-Being and Professional Performance in Nursing Students. *Nurse Education Today* 31, 8 (2011), 855–860.
- [76] Soujanya Poria, Erik Cambria, Rajiv Bajpai, and Amir Hussain. 2017. A Review of Affective Computing: From Unimodal Analysis to Multimodal Fusion. *Information Fusion* 37 (2017), 98–125.
- [77] J Donald Ragsdale. 1976. Relationships between hesitation phenomena, anxiety, and self-control in a normal communication situation. *Language and Speech* 19, 3 (1976), 257–265.
- [78] Hedieh Ranjbartabar and Deborah Richards. 2016. A Virtual Emotional Freedom Therapy Practitioner. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1471–1473.
- [79] Peter AM Ruijten, Cees JH Midden, and Jaap Ham. 2016. Ambiguous Agents: The Influence of Consistency of an Artificial Agent’s Social Cues on Emotion Recognition, Recall, and Persuasiveness. *International Journal of Human-Computer Interaction* 32, 9 (2016), 734–744.
- [80] Timothy Rumbell, John Barnden, Susan Denham, and Thomas Wennekers. 2012. Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Autonomous Agents and Multi-Agent Systems* 25, 1 (2012), 1–45.
- [81] Stuart Russell and Peter Norvig. 1995. A modern approach. *Artificial Intelligence. Prentice-Hall, Englewood Cliffs* 25 (1995), 27.
- [82] Nuray Sakalli-Ugurlu, Selin Salman, and Sinem Turgut. 2010. Predictors of Turkish Women’s and Men’s Attitudes toward Sexual Harassment: Ambivalent Sexism, and Ambivalence toward Men. *Sex Roles* 63, 11-12 (2010), 871–881.
- [83] Peter Salovey and Daisy Grewal. 2005. The Science of Emotional Intelligence. *Current Directions in Psychological Science* 14, 6 (2005), 281–285.
- [84] Peter Salovey and John D Mayer. 1990. Emotional intelligence. *Imagination, Cognition and Personality* 9, 3 (1990), 185–211.
- [85] Carla Sharp, Salome Vanwoerden, Y Van Baardewijk, JL Tackett, and H Stegge. 2015. Callous-unemotional traits are associated with deficits in recognizing complex emotions in preadolescent children. *Journal of personality disorders* 29, 3 (2015), 347–359.
- [86] Joanna Stern. 2017. Alexa, siri, cortana: The problem with all-female digital assistants. <https://www.wsj.com/articles/alexa-siri-cortana-the-problem-with-all-female-digital-assistants-1487709068> Accessed on February 2017.
- [87] Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing thought: improving robot readability with animation principles. In *Proceedings of the sSixth International Conference on Human-Robot Interaction*. ACM, 69–76.
- [88] Maya Tamir. 2016. Why do people regulate their emotions? A taxonomy of motives in emotion regulation. *Personality and Social Psychology Review* 20, 3 (2016), 199–222.
- [89] Jianhua Tao and Tieniu Tan. 2005. Affective Computing: A Review. In *International Conference on Affective Computing and intelligent interaction*. Springer, 981–995.
- [90] Christiana Tsiouri, Maher Ben Moussa, João Quintas, Ben Loke, Inge Jochem, Joana Albuquerque Lopes, and Dimitri Konstantas. 2016. A Virtual Assistive Companion for Older Adults: Design Implications for a Real-World Application. In *Proc. SAI Intell. Syst. Conf.* 556–566.
- [91] Gerben A Van Kleef, Helma van den Berg, and Marc W Heerdink. 2015. The Persuasive Power of Emotions: Effects of Emotional Expressions on Attitude Formation and Change. *Journal of Applied Psychology* 100, 4 (2015), 1124.
- [92] Thuri Vogt, Elisabeth André, and Nikolaus Bee. 2008. EmoVoice: A Framework for Online Recognition of Emotions from Voice. In *International Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems*. Springer, 188–199.
- [93] Elaine F Walker and Jayne E Stake. 1978. Changes in preferences for male and female counselors. *Journal of Consulting and Clinical Psychology* 46, 5 (1978), 1153.
- [94] Jing Wang, Ronald J Iannotti, and Tonja R Nansel. 2009. School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *Journal of Adolescent Health* 45, 4 (2009), 368–375.
- [95] Jerry S Wiggins. 1979. A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology* 37, 3 (1979), 395.
- [96] Melissa J Williams and Larissa Z Tiedens. 2016. The subtle suspension of backlash: A meta-analysis of penalties for women’s implicit and explicit dominance behavior.
- [97] Lauren Wispé. 1987. History of the Concept of Empathy. *Empathy and Its Development* (1987), 17–37.
- [98] Steven B Wolff, Anthony T Pescosolido, and Vanessa Urch Druskat. 2002. Emotional Intelligence as the Basis of Leadership Emergence in Self-Managing Teams. *The Leadership Quarterly* 13, 5 (2002), 505–522.
- [99] Michael Wooldridge and Nicholas R Jennings. 1995. Intelligent Agents: Theory and practice. *The Knowledge Engineering Review* 10, 02 (1995), 115–152.
- [100] Yang Yang, Xiaojuan Ma, and Pascale Fung. 2017. Perceived Emotional Intelligence in Virtual Agents. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM.