

Communicating with Machines: Conversational Agents with Personality and the Role of Extraversion

Rangina Ahmad
TU Braunschweig
rangina.ahmad@tu-bs.de

Dominik Siemon
TU Braunschweig
d.siemon@tu-bs.de

Susanne Robra-Bissantz
TU Braunschweig
s.robra-bissantz@tu-bs.de

Abstract

Communication with conversational agents (CA) has become increasingly important. It therefore is crucial to understand how individuals perceive interaction with CAs and how the personality of both the CA and the human can affect the interaction experience. As personality differences are manifested in language cues, we investigate whether different language style manifestations of extraversion lead to a more anthropomorphized perception (specifically perceived humanness and social presence) of the personality bots. We examine, whether individuals rate communication satisfaction of a CA similar to their own personality as higher (law of attraction). The results of our experiment indicate that highly extraverted CAs are generally better received in terms of social presence and communication satisfaction. Further, incorporating personality into CAs increases perceived humanness. Although no significant effects could be found in regard to the law of attraction, interesting findings about ambiverts could be made. The outcomes of the experiment contribute towards designing personality-adaptive CAs.

1. Introduction

In the past years, progress has been made in the field of AI in recognizing and mimicking human activities [22, 29]. Conversational Agents (CA) are to a certain degree capable of simulating such human behavior, by interacting with their users via natural language – in written or spoken form [22, 35]. With the introduction of Google Assistant, Apple's Siri or Amazon's Alexa, CAs have been made accessible to a wide range of users [35]. CAs have also been incorporated in many companies, specifically on their websites and messenger platforms, for example to assist customers during the sales process [31]. Another context, in which CAs play an increasingly important role is in health and medical care, supporting consumers with mental health challenges, or assisting patients and elderly individuals in their living

environments [32]. With all these areas of application, the design of CAs and specifically their effect on human-computer interaction (HCI) has become even more important. In prolonging casual conversations, the long-term goal is to take steps toward the AI dream of general intelligence, that is, CAs that are able to provide casual opinions or therapeutic responses and have “a lot of personality” [3, p.3]. However, handling open conversation effectively and providing the machine with the ability to converse with humans in a natural way and in a manner that the user is satisfied or even enjoys the interaction, is still a major challenge [22, 35, 59].

In order to have a better understanding of the nature and quality of human-machine interactions, researchers have drawn from a variety of other disciplines, such as psychology and sociology. In this context, many studies have confirmed that findings in human-human interaction (HHI) can be applied to the interaction between machines and humans [36, 39, 41, 47]. For instance, people treat a computer system as if it were a human and thus project a certain level of anthropomorphism upon the machine – this perceived humanness has been an important aspect for social interactions [14, 37, 40, 52]. Also examined by many researchers is the construct of social presence, which describes the sense of connection that a user feels with their IT communication partner [12, 19, 52]. As a matter of fact, humans will respond to computers and CAs using the same elements of social interactions that they employ in HHI [37, 41]. In particular, researchers have shown that people respond to computer systems in similar ways to how they would respond to a human, for instance by attributing certain personality traits to computer partners [41, 52]. A specific set of characteristics is on the one hand believed to explain the way people respond to others in social settings [47, 57], and provides on the other hand an explanation why it influences the quality of interactions [13, 47]. As for the question, which personality traits in specific are more desirable in CAs, there is no definite answer to it. In fact, more and more researchers conclude that CAs should not have a static personality, but should

ideally adapt to the personality of the user [1, 2, 22, 53, 55, 64]. As part of initial steps towards our research goal of designing personality-adaptive CAs, we investigate existing theories, specifically social presence and perceived humanness in connection with the personality trait extraversion. The extraversion/introversion dichotomy has specifically received the most attention from researchers, as the underlying components of extraversion have been well-established to date across various methodologies [7, 33]. Further, since incorporating different personality-based language styles increase social presence [42] and extraverted individuals are more likely to anthropomorphize bots [47, 49], we pose the following research question (RQ):

RQ1: Are social presence and perceived humanness increased when incorporating an extraverted language style in human-machine conversation?

Another step towards a better understanding of how to increase interaction quality provides the law of attraction. According to this theory, humans prefer machines with a similar personality to their own [39, 41]. On the one hand, studies confirm the law of attraction [39, 41], on the other hand however, there are some findings indicating people prefer interaction based on opposing personality traits [27]. Since in this case concepts of HHI cannot be clearly transferred to HCI, and since there is a larger body of research speaking for the law of attraction, we base our second RQ on this theory and investigate the following question:

RQ2: Are extraverts perceiving higher communication satisfaction when interacting with an extraverted CA?

Though researchers have given great thought to the effects of different personality dimensions on HCI, yet, only a few studies have dedicated their research on personality – extraversion in particular – and their effect on perceived humanness and social presence. Further, as there are still disagreements regarding the law of attraction and its opposing theory, we aim to study whether different language style manifestations of extraversion (i.e. high, medium, low) have an effect on users' perceived communication satisfaction.

The paper is structured as follows: First, we provide relevant theoretical background on personality-based language cues, social presence and perceived humanness as well as on the law of attraction. We derive our hypotheses, which are then tested in the

framework of our conducted experiment. We present the results and discuss our research questions.

2. Theory & Hypotheses

Personality has been an essential topic in the literature of HCI. Personality differences are manifested in language and can thus be incorporated into intelligent systems such as CAs by using specific language cues that reflect a certain personality dimension. CAs can have (or simulate) a personality that can easily adapt according to a user's personality to give the user the best possible interaction. In research, the law of attraction is often used to determine which personality types and factors are preferred. However, an important prerequisite for this is the general perception of a CA by its user and specifically their perceived humanness and social presence

2.1. Personality & Language Cues

A well accepted theory of psychology is that human language reflects the emotional state and personality, based on the frequency with which certain categories of words are used as well as the variations in word usage [7, 20, 63]. Personality is described as a psychological construct that comprises someone's behaviors, emotions and cognitions derived from both biological and social factors [23, 47]. Theories of psychology further assert that personality traits in particular can be used to predict human emotions, cognitions and behaviors [42, 47]. In order to measure an individual's personality, a widely used classification of personality – the Big Five model – has been applied in research [34]. For a comprehensive assessment of individuals, the following five fundamental traits or dimensions have been defined and derived through factorial studies: Conscientiousness, openness, neuroticism, agreeableness and extraversion which refers to the extent to which people enjoy company and seek excitement and stimulation [10]. This specific set of characteristics is believed to explain the way people respond to others in social settings, [47, 57] and also provides an explanation as to why it influences the quality of interactions between people [13, 42, 47]. Verbal interactions by means of language can therefore be useful for capturing lower-level personality processes, since language is more closely associated with objective behavioral outcomes than traditional personality measures [7]. Utterances for instance - conveying a great deal of information about the speaker - comprise cues to the individual's personality traits [33]. Both spoken language as well as written

language is unique from person to person, as humans express themselves verbally with their own distinctive styles, even though the content of a message may be the same [43]. By discovering correlations between a range of linguistic variables and personality traits across a wide range of linguistic levels, psychologists have documented the existence of such cues [33].

Although research has found language markers for all Big Five traits, the extraversion/ introversion dichotomy has specifically received the most attention from researchers, as the underlying components of extraversion have been well-established to date across various methodologies [7, 33]. Most of the personalities can be measured somewhere between two extremities, since extraversion-introversion levels are part of a single, continuous dimension of personality [56]. Table 1 shows a small overview of some of the identified language cues for extraversion and introversion, based on various studies by Scherer (1979) [50], Furnham (1990) [15], Pennebaker and King (1999) [43], Gill and Oberlander (2002) [18] and Mairesse et al. (2007) [33].

2.2. Social Presence & Perceived Humanness

Social presence refers to the subjective perceptions of a medium that is used for communication, i.e. the sense of connection that a user feels with their communication partner [52]. It measures the extent to which communication media has the ability to transmit social cues, such as transmitting sociable, personal or intimate aspects [40, 52]. While early research focused on computer-mediated communication between humans, recent studies state that social presence is also

applicable to HCI and consequently to CAs [12, 37]. According to the ‘computers are social actors’ (CASA) paradigm [37, 40, 41], humans will respond to computers and CAs using the same elements of social interactions that they employ in HHI [41]. In a variety of studies, researchers have shown that people respond to computer systems in similar ways to how they would respond to a human – for instance by attributing specific personality traits to computer partners [37, 52]. Relative to introverts, extraverts generally engage in more social activity, experience greater positive affect and well-being, and are reactive to external stimulation [15, 33, 50]. Further, an extraverted person will be relatively gregarious and generally energized by external stimulus or people [16, 58]. An introvert on the contrary, will be comparatively less sociable and more introspective [41, 54, 58]. These tendencies indicate that humans may apply these social rules equally to HCI.

The construct of social presence can be used to measure a user’s perception of a CA’s social characteristics. By incorporating a language style, that reflects a certain personality dimension, social presence can be increased [44]. A high degree of social presence indicates enhanced trusting beliefs and perceptions of enjoyment in the interaction with a CA [45]. We therefore hypothesize:

H1a: Social presence (SP) is perceived higher for extraverted than for introverted bots.

Perceived humanness is another relevant factor in the context of HCI. When interacting with an intelligent technological artifact such as a CA, the

Table 1. Language cues for extraversion and introversion

Level	Introvert	Extravert
Conversational Behavior	Listen, less back-channel behavior	Initiate conversation, more back-channel behavior
Style	Formal	Informal
Syntax	Many nouns, adjectives, elaborated constructions, many words per sentence, many articles and negations	Many verbs, adverbs, pronouns (implicit), few words per sentence, few articles, few negations
Topic selection	Self-focused, problem talk, dissatisfaction, single topic, few semantic errors	Pleasure talk, agreement, compliment, many topics, many semantic errors
Speech	Slow speech rate, Many unfilled pauses, long response latency, quiet, low voice quality, low frequency variability	High speech rate, few unfilled pauses, short response latency, loud, high voice quality, high frequency variability
Lexicon	Rich, high diversity, many exclusive and inclusive words, few social words, few positive emotion words, many negative emotion words	Poor, low diversity, few exclusive and inclusive words, many social words, many positive emotion words, few negative emotion words

degree to which it is perceived as human has a positive impact on aspects like trust [51] and service encounter satisfaction [19]. Furthermore, a high degree of perceived humanness increases the degree of perceived competency in the CA [3], which in turn can improve the overall interaction. Perceived humanness is also related to the willingness to disclose personal information and the degree to which users feel more comfortable interacting with a CA [38]. Since people project a certain level of anthropomorphism upon the computer, they treat the computer system as if it were a human [37, 40, 52]. Specifically individuals who scored high in extraversion have been found to be more likely to anthropomorphize a bot [47, 49]. This leads us to the following hypothesis regarding perceived humanness:

H1b: Perceived humanness (PH) is perceived higher for extraverted than for introverted bots.

2.3. Law of Attraction

Since personality may be one of the most important factors in HCI among others (specifically when designing social CAs), studies have tried to define social CAs in order to determine their effect in human-machine interactions. A rich history of research suggests that HCI largely parallel HHI [37, 39, 40, 46, 48], which is why researchers have examined machine personality based on human personality characteristics [39, 46, 48]. The law of attraction (also called similarity-attraction theory) as one of the HHI theories that can be transferred to HCI, posits that humans prefer machines that have a similar personality to their own and are thus more comfortable interacting with them than with those CAs with an opposing personality [39, 41]. Users can better assess their counterparts when using similar personality traits, and information has usually been rated as better and more trustworthy [64].

As for the trait extraversion, Moon and Nass (1996) [36] found that dominant people prefer to interact with an equally dominant chatbot. The same applies to people with a submissive behavior [36]. Concerning verbal communication in HHI, it has been shown that speaker charisma for example strongly correlates with extraversion [33]. Relative to their introverted counterparts, extraverts tend to talk more, with fewer pauses and hesitations, have shorter silences, a higher verbal output and a less formal language, while introverts use a broader vocabulary [15, 18, 43, 50]. Research also showed that conversations between extraverts are more expansive and characterized by a wider range of topics whereas a conversation between two introverts are more serious and have a greater topic

focus (i.e., discussing one topic in depth) [15]. Extraverts exert a more imprecise and “looser” style with reduced concreteness, whereas introverts exhibit a more analytic, careful, precise and focused style [18]. Extraverts also use more positive emotion words and show more agreements and compliments than introverts [43]. We therefore make the following two hypotheses:

H2a: Extraverts assess their perceived communication satisfaction for the extraverted bot higher as for the introverted bot.

H2b: Introverts assess their perceived communication satisfaction for the introverted bot higher as for the extraverted bot.

Since not every individual falls into one of the two extremes, but rather somewhere in the middle of the personality spectrum, a third group called ambiverts has been identified by psychologists [56]. As a solid mix of both extraversion and introversion, ambiverts are balanced between the extremes [5]. As ambiverts express traits from both personality styles and sometimes neither is dominant [21, 56], ambiverts are often more flexible in their communication and interaction as they draw from a broader range of communication options [14]. For instance, they engage in a flexible pattern of talking and listening [5]. However, while ambiverts are able to adapt more easily to the demands of a specific situation, they sometimes have trouble determining which side of their personality to apply [56]. Based on these findings, we hypothesize as follows:

H2c: Ambiverts do not assess communication satisfaction differently for extraverted or introverted bots.

3. Method

In order to address our research question and test our hypotheses, we conducted an online experiment. We chose an experimental study to maximize internal validity [30] and to be able to validate if any observed differences in communication satisfaction, perceived humanness and social presence, i.e. the dependent variables are caused by our manipulation, i.e. the personality of our CAs (independent variable) [8]. Our experiment follows a within-subject design, in which participants experience both conditions. The experiment took course over the span of two month and had on average a completion time of 20 - 25 minutes.

3.1. Sample

A relatively large sample size of participants was obtained to ensure more reliable results and greater significance. 478 native English speaker or people with English as their second language were recruited from the crowdsourcing platform Mechanical Turk (mTurk) as well as via personal networks. While the participants on mTurk were paid for the task, the people from our personal network were not compensated for their participation. Subjects of both sources did not differ regarding any relevant characteristics, such as their demographic data or personality traits. Of the 478 participants, we eliminated the data of 113 test persons who cancelled the experiment in advance. Another 102 subjects failed the attention checks built into the study or did not complete the personality test, leaving 263 participants (177 males, 84 females, 2 other). The age of the subjects ranged from 17 to 74 years, with an average age of 32.9 years.

3.2. Procedure

In the first step of the experiment, the subjects were asked to conduct an online personality test based on Johnson's (2014) [28] 120-item IPIP NEO-PI-R. Their results were values between 0-120, with 0 being a low score and 120 a high score in a specific personality dimension. The unique identification numbers the participants received after finishing their personality tests were also used throughout the second part of the experiment in order to provide anonymity between the experimenters and the subjects. We chose a within-subject design, where the participants were exposed to both levels of treatment one after the other [9], i.e. the subjects were randomly assigned to LimeSurvey (an online survey tool), where they either watched a video of a conversation between an extraverted Chatbot (ExtraBot) and a fictional human first and a conversation of the introverted Chatbot (IntroBot) second or vice versa. This way we ensured that individual differences were not distorting the results, since every subject acted as their own control. This reduced the chance of confounding factors. The order of the two conditions was hence distributed randomly, and the dependent variable was measured after each condition by means of a subsequent survey. Every participant was provided with the exact same sets of information for the experiment [11].

The videos lasted about 3 minutes, skipping was not allowed, and a control question was added at the end of the video. The participants' task was to put themselves in the shoes of the human conversation

partners and to closely observe the dialogues with Raffi the CA. Subsequently, the subjects completed a survey including the established measuring construct communication satisfaction (CS) by Hecht (1978) [24], in order to test H2a, H2b and H2c. Originally intended for HHI, we transferred this construct to a HCI context, since the inventory shows a high degree of reliability and validity when measuring communication satisfaction with "actual and recalled conversations with another perceived to be a friend, acquaintance, or stranger" [17, p. 253]. The construct consists of 19 items, and as suggested in the study, we used a 7-point Likert scale. However, we adapted the phrasing of the items accordingly to our CA. For instance, we changed the wording of the original item "The other person expressed a lot of interest in what I had to say" to "Raffi expressed a lot of interest in what I had to say".

In order to check hypotheses H1a and H1b, the survey included the measures social presence (SP) and perceived humanness (PH). We used the perceived humanness measure (6 items) by Gill and Oberlander (2002) [17], which is measured on a 9-point Likert scale (1 = low, 9 = high) and social presence (5 items) by Gefen and Straub (2004) [17], which was measured on a 7-point Likert scale (1 = low, 7 = high). The survey further covered demographic questions (gender, age, language) and an open question (Do you think that the concept of a personality adaptive CA can enhance a person's interaction experience?). To check if our manipulation was successful and the participants perceived the ExtraBot as more extraverted as the IntroBot, we asked the participants to indicate the extent to which the attributes sociable, talkative, active, impulsive, outgoing, shy, reticent, passive, deliberate, reserved apply to the CAs [4]. While the first five items reflect high extraversion, the last five attributes stand for low extraversion [4]. The measure is summarized as the construct extraversion (EX).

3.3. Conversational Agent Design

We used the conversational design tool Botsociety [6] to visualize our CAs. The two pre-defined dialogue structures are communications between the CA Raffi and the humans Jamie and Francis, respectively. While Jamie and the ExtraBot are intended to take on an extraverted personality, Francis and the IntroBot are both of introverted nature. We achieved this by adapting their respective language style according to the personality dimensions, specifically by using the language cues for extraversion and introversion listed in Table 1. In order to find out whether the language styles used in the conversations truly reflected the dimensions extraversion and introversion, we used IBM Watson's Personality Insights tool [26] for

verification purposes. Based on text that is being analyzed, the personality mining service returns different percentiles for the Big Five dimensions. Percentiles are defined as scores that compare one person to a broader population [26]. While the ExtraBot dialogue received a score of 83%, meaning that the ExtraBot is more extraverted than 83% of the people in the population, the IntroBot had a percentile of 36%, thus scoring low in extraversion (and high in introversion). With these results, we incorporated the two language styles into our CA design. The topics of the dialogues were everyday conversations, such as weekend plans, music, books and travel. However, compared to the ExtraBot, the IntroBot sticks mainly to two topics (travel and books) and has a rather rich vocabulary throughout the whole dialogue by using many words per sentence. The IntroBot also uses fewer emotional words and has fewer semantic errors (“At least it's Friday! How was your day?”) compared to the ExtraBot. The ExtraBot on the other hand talks about many topics in a short amount of time, uses a rather informal language (e.g. “what r u up to?”, “...cause TGIF!”), compliments and uses many positive emotion words (“Sounds amazing!” “Have fun at the party!”) and few words per sentence (“Nope. Locals as well.”, “Told ya!”).

The videos of the complete conversations can be watched at the following links: <https://youtu.be/B1N7XwcdCE0>, <https://youtu.be/d26eKdHBKeQ>. Figure 1 shows a snippet of the two conversations between the ExtraBot and Jamie (left) and the IntroBot and Francis (right).



Figure 1. Mockups of the ExtraBot (left) and IntroBot (right) Conversation

3.4. Results

We used confirmatory factor analysis to test our measures CS, PH, SP and EX (manipulation check). After dropping three items of CS and three items of EX (manipulation check), all items had significant positive factor loadings greater than .600. Furthermore, we calculated composite reliability (CR) and Cronbach's alpha, both indicating reliable factors (see Table 2) [60].

All analyses were carried out using the statistical computing software RStudio (Version 1.2.5033). Due to the use of ordinal scales [62], a non-normality assumption and a within-subject design (paired sample), we chose the nonparametric Wilcoxon signed-rank test to test for significant differences (CI = .95). Table 2 provides the descriptive statistics, Cronbach's alphas and the composite reliability of the measures.

Table 2. Descriptive statistics, Cronbach's alpha and composite reliability

Measure	M _{EB}	SD _{EB}	M _{IB}	SD _{IB}	α	CR
CS	5.25	0.98	5.08	0.94	.90	.89
PH	7.22	1.35	7.08	1.48	.91	.91
SP	5.50	1.05	5.30	1.19	.89	.89
EX	4.68	0.86	4.11	0.89	.81	.73

After the experiment, the participants indicated the extent to which the attributes concerning extraversion apply to the ExtraBot and IntroBot. The conducted Wilcoxon test revealed, that there is a significant difference ($W = 46374$, $p < .001$) in the rated attributes of extraversion between the ExtraBot and the IntroBot i.e. the ExtraBot was perceived as more extraverted than the IntroBot.

Since our manipulation check was successful, we continued testing our hypotheses. Social presence was rated higher for the ExtraBot ($M = 5.50$) than for the IntroBot ($M = 5.30$), which is a significant difference ($W = 37826$, $p = .048$), supporting H1a. perceived humanness was rated slightly higher for the ExtraBot ($M = 7.22$) than for the IntroBot ($M = 7.08$). However, the difference is not significant ($W = 36981$, $p = .169$), which is why H1b cannot be supported.

In order to test H2a, H2b and H2c, the population was divided into three different groups, representing participants with a low extraversion (introverts), medium extraversion (ambiverts) and high extraversion (extraverts). Clustering was performed by means of the R package *Ckmeans.1d.dp*, a procedure for optimal k-means clustering in one dimension [61]. This resulted in three groups with $n = 58$ for introverts (Cluster center = 55.20, Min = 30, Max = 66), $n = 114$ for ambiverts (Cluster center = 77.81, Min = 67, Max = 85) and $n = 91$ for extraverts (Cluster

center = 93.20, Min = 86, Max = 111). Table 3 shows the descriptive statistic of each cluster, for communication satisfaction and the results of the Wilcoxon tests.

Table 3. Communication satisfaction by clusters

	Introverts	Ambiverts	Extraverts
n	58	114	91
Communication satisfaction			
M _{EB}	5.40	5.03	5.42
M _{IB}	5.30	4.84	5.20
Wilcoxon test	W = 4380 p = 0.501	W = 7152 p = 0.189	W = 2036 p = 0.113

The results of the Wilcoxon tests show no significant difference in communication satisfaction between the ExtraBot and IntroBot within the clusters. Therefore, only H2c is supported, whereas we reject H2a and H2b.

4. Discussion

The mean values of the PH measure (PH_{EB} = 7.22; PH_{IB} = 7.08) are above the mean of the scale (4.50) for both the ExtraBot and the IntroBot. These substantially increased values demonstrate that the incorporation of language cues and the associated depiction of a personality has led to a human perception of the CAs, even though there is no significant difference between the two CAs. The SP scale also clearly shows an increased value for both CAs (SP_{EB} = 5.50; SP_{IB} = 5.30; Mean of the scale = 3.50). The Wilcoxon test furthermore reveals a significant difference between the ExtraBot and the IntroBot for SP. SP was rated higher for the ExtraBot than for the IntroBot, indicating that the ExtraBot's ability to transmit social cues, its personality and sociability with the incorporation of language cues was better than the IntroBot. Concerning RQ1, both SP and PH were increased by personality differences manifested in language use. However, while an extraverted language style as opposing to an introverted language use led to a significantly higher SP, the same cannot be said for PH. The subjects did not perceive the ExtraBot as significantly more human than the introverted CA. Nevertheless, the results show that the use of personality-based language styles in HCI can achieve an increased SP as well as PH. These theories in turn are fundamental aspects for an enhanced interaction quality.

Clustering the participants according to their level of extraversion, resulted in a large number of ambiverts (n = 114), followed by the second biggest group of extraverts (n = 91) and the smallest group consisting of

introverts (n = 58). The results indicate that all three groups rated CS higher for the ExtraBot, even though no significant difference was found. So while hypotheses H1a and H1b cannot be supported, it still can be said that overall the participants rated CS higher for the ExtraBot (M = 5.25) than for the IntroBot (M = 5.08), which is a significant difference (W = 38057, p = .046). To our surprise, introverts rated the CS of the ExtraBot similar high as the extraverts, thus not confirming the law of attraction. A possible explanation for this may be that the introverted participants were more drawn to an extraverted CA, because the ExtraBot's SP was significantly higher than the IntroBot's SP. And as studies have shown, the higher the degree of social presence, the more trust beliefs and perceptions of enjoyment in the interaction with a CA a user has. As assumed for the ambiverts, the results show that they rated CS in both CAs as the lowest. H2c thus can be supported, as they expressed no clear preference for any of the two bots.

In order to find a clearer answer for RQ2, we took a closer look at the dataset, particularly the data of participants who scored either very high or very low in extraversion. The tendencies towards their bot preferences were mixed: We found extraverts, who clearly rated the CS for the ExtraBot as much higher as for the IntroBot (confirming the law of attraction), but then again there were also participants high in extraversion, who indicated that they perceived a higher CS for the IntroBot. The same could be observed vice versa for introverts. The rather high number of outliers, who preferred the CA with the opposing personality trait explain why there could not be found a significant effect on CS. We further believe, a direct interaction between the participants and the CAs - rather than letting the subjects evaluate the conversation of two fictional characters and assuming all participants were able to empathize with Jamie and Francis - may have had a significant effect on CS. So the question, whether extraverts perceive higher CS when interacting with an extraverted CA cannot be answered with a definite yes.

Though 3 out of 5 hypotheses were rejected - in regard to our research project towards designing personality-adaptive CAs, we take the following aspects from this experiment: First, incorporating personality that is manifested in a specific language style (in this case notably extraversion), increases a bot's social presence. Second, using language cues that are specific for a certain personality dimension (whether high or low extraversion), lead to the CA being perceived more as a human. Third, as a high number of the population is believed to be ambiverted, rather than extremely introverted or extraverted [56]

(matching our dataset), further research has to be done on how to enhance ambiverts' interaction quality. In fact, since ambiverts seem to have no specific preferences concerning their communication style and rather adapt according to a given situation, it is the more important that a CA itself should be able to switch and adapt to the preferred communication style of its user. Fourth, the results of our experiment clearly demonstrated that the law of attraction cannot always be confirmed. The number of participants who showed a preference for a CA with an opposing personality trait were too high to ignore and label them just as "a few outliers". This again argues for the fact that CA personality should not be designed statically but dynamically, so that it easily adapts to user personality - regardless of whether the user is highly extraverted, introverted or has any other dominant Big Five personality trait. It can be concluded that a "one-size-fits-all"- design for CAs does not necessarily enhance interaction quality, since each user has unique personality traits, abilities, perceptual preferences, experiences, etc., that directly affect every interaction process. Thus, communicating with machines should be personalized.

5. Conclusion

As an initial step towards our research goal of designing personality-adaptive CAs, we investigated existing theories, such as social presence, perceived humanness and the law of attraction in connection with the personality trait extraversion. Our experiment demonstrated that incorporating personality into a bot, makes the CA being perceived as more human. As opposing to the introverted bot, the extraverted CA in particular was better received in terms of social presence. Though communication satisfaction of the extraverted CA was slightly rated higher, no significant effects could be found in regard to the law of attraction. Interesting findings about ambiverts could be made, however more research is required. Our findings contribute to CA design and provide insights for researchers who want to follow up with research on personality in CAs.

6. References

- [1] Ahmad, R., D. Siemon, D. Fernau, and S. Robra-Bissantz, "Introducing" Raffi": A Personality Adaptive Conversational Agent.", *PACIS 2020*, 28.
- [2] Ahmad, R., D. Siemon, and S. Robra-Bissantz, "ExtraBot vs IntroBot: The Influence of Linguistic Cues on Communication Satisfaction", *AMCIS 2020*.
- [3] Araujo, T., "Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions", *Computers in Human Behavior* 85, 2018, pp. 183–189.
- [4] Back, M.D., S.C. Schmukle, and B. Egloff, "Predicting actual behavior from the explicit and implicit self-concept of personality.", *Journal of personality and social psychology* 97(3), 2009, pp. 533.
- [5] Bernstein, E., "Not introvert, nor extrovert: The adaptable ambivert.", *Wall Street Journal*, 2015, pp. 1–3.
- [6] Botsociety, "Design, preview and prototype your next chatbot or voice assistant", 2020. <https://botsociety.io>
- [7] Boyd, R.L., and J.W. Pennebaker, "Language-based personality: a new approach to personality in a digital world", *Current Opinion in Behavioral Sciences* 18, 2017, pp. 63–68.
- [8] Brewer, M.B., and W.D. Crano, "Research design and issues of validity", *Handbook of research methods in social and personality psychology*, 2000, pp. 3–16.
- [9] Charness, G., U. Gneezy, and M.A. Kuhn, "Experimental methods: Between-subject and within-subject design", *Journal of Economic Behavior & Organization* 81(1), 2012, pp. 1–8.
- [10] Costa, P.T., and R.R. McCrae, "The revised neo personality inventory (neo-pi-r)", *The SAGE handbook of personality theory and assessment* 2(2), 2008, pp. 179–198.
- [11] Dennis, A.R., and J.S. Valacich, "Conducting experimental research in information systems", *Communications of the association for information systems* 7(1), 2001, pp. 5.
- [12] Diederich, S., M. Janssen-Müller, A.B. Brendel, and S. Morana, "Emulating Empathetic Behavior in Online Service Encounters with Sentiment-Adaptive Responses: Insights from an Experiment with a Conversational Agent", *ICIS 2019 Proceedings*, 2019.
- [13] Driskell, J.E., G.F. Goodwin, E. Salas, and P.G. O'Shea, "What makes a good team player? Personality and team effectiveness.", *Group Dynamics: Theory, Research, and Practice* 10(4), 2006, pp. 249.
- [14] Fink, J., "Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction", *Social Robotics*, Springer (2012), 199–208.
- [15] Furnham, A., "Faking personality questionnaires: Fabricating different profiles for different purposes", *Current psychology* 9(1), 1990, pp. 46–55.
- [16] Furnham, A., and C.R. Brewin, "Personality and happiness", *Personality and individual differences* 11(10), 1990, pp. 1093–1096.
- [17] Gefen, D., and D.W. Straub, "Consumer trust in B2C e-Commerce and the importance of social presence:

- experiments in e-Products and e-Services”, *Omega* 32(6), 2004, pp. 407–424.
- [18] Gill, A.J., and J. Oberlander, “Taking care of the linguistic features of extraversion”, *Proceedings of the Annual Meeting of the Cognitive Science Society*, (2002).
- [19] Gnewuch, U., S. Morana, and A. Maedche, “Towards Designing Cooperative and Social Conversational Agents for Customer Service”, *ICIS*, (2017).
- [20] Golbeck, J., C. Robles, M. Edmondson, and K. Turner, “Predicting personality from twitter”, *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, IEEE (2011), 149–156.
- [21] Grant, A.M., “Rethinking the extraverted sales ideal: The ambivert advantage”, *Psychological science* 24(6), 2013, pp. 1024–1030.
- [22] Grudin, J., and R. Jacques, “Chatbots, Humbots, and the Quest for Artificial General Intelligence”, *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, ACM Press (2019), 1–11.
- [23] Hall, C.S., G. Lindzey, and J.B. Campbell, *Theories of personality*, Wiley New York, 1957.
- [24] Hecht, M.L., “The Conceptualization and Measurement of Interpersonal Communication Satisfaction”, *Human Communication Research* 4(3), 1978, pp. 253–264.
- [25] Holtgraves, T., and T.-L. Han, “A procedure for studying online conversational processing using a chat bot”, *Behavior research methods* 39(1), 2007, pp. 156–163.
- [26] IBM Watson PI, “IBM Watson Personality Insights”, 2020. <https://personality-insights-demo.ng.bluemix.net/>
- [27] Isbister, K., and C. Nass, “Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics”, *International Journal of Human-Computer Studies* 53(2), 2000, pp. 251–267.
- [28] Johnson, J.A., “Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120”, *Journal of Research in Personality* 51, 2014, pp. 78–89.
- [29] Kaplan, A., and M. Haenlein, “Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence”, *Business Horizons* 62(1), 2019, pp. 15–25.
- [30] Karahanna, E., I. Benbasat, R. Bapna, and A. Rai, “Editor’s comments: opportunities and challenges for different types of online experiments”, *MIS Quarterly* 42(4), 2018, pp. iii–x.
- [31] Knijnenburg, B.P., and M.C. Willemsen, “Inferring capabilities of intelligent agents from their external traits”, *ACM Transactions on Interactive Intelligent Systems (TiiS)* 6(4), 2016, pp. 28.
- [32] Laranjo, L., A.G. Dunn, H.L. Tong, et al., “Conversational agents in healthcare: a systematic review”, *Journal of the American Medical Informatics Association* 25(9), 2018, pp. 1248–1258.
- [33] Mairesse, F., M.A. Walker, M.R. Mehl, and R.K. Moore, “Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text”, *Journal of Artificial Intelligence Research* 30, 2007, pp. 457–500.
- [34] McCrae, R.R., and O.P. John, “An introduction to the five-factor model and its applications”, *Journal of personality* 60(2), 1992, pp. 175–215.
- [35] McTear, M., Z. Callejas, and D. Griol, *The conversational interface: Talking to smart devices*, Springer, 2016.
- [36] Moon, Y., and C. Nass, “How ‘real’ are computer personalities? Psychological responses to personality types in human-computer interaction”, *Communication research* 23(6), 1996, pp. 651–674.
- [37] Nass, C., and Y. Moon, “Machines and Mindlessness: Social Responses to Computers”, *Journal of Social Issues* 56(1), 2000, pp. 81–103.
- [38] Nass, C., and Y. Moon, “Machines and mindlessness: Social responses to computers”, *Journal of social issues* 56(1), 2000, pp. 81–103.
- [39] Nass, C., Y. Moon, B.J. Fogg, B. Reeves, and D.C. Dryer, “Can computer personalities be human personalities?”, *International Journal of Human-Computer Studies* 43(2), 1995, pp. 223–239.
- [40] Nass, C., J. Steuer, and E.R. Tauber, “Computers are social actors”, *Proceedings of the SIGCHI conference on Human factors in computing systems*, (1994), 72–78.
- [41] Park, E., D. Jin, and A.P. del Pobil, “The law of attraction in human-robot interaction”, *International Journal of Advanced Robotic Systems* 9(2), 2012, pp. 35.
- [42] Peeters, M.A., H.F. Van Tuijl, C.G. Rutte, and I.M. Reymen, “Personality and team performance: a meta-analysis”, *European Journal of Personality: Published for the European Association of Personality Psychology* 20(5), 2006, pp. 377–396.
- [43] Pennebaker, J.W., and L.A. King, “Linguistic styles: Language use as an individual difference.”, *Journal of personality and social psychology* 77(6), 1999, pp. 1296.
- [44] Qiu, L., and I. Benbasat, “Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems”, *Journal of Management Information Systems* 25(4), 2009, pp. 145–182.

- [45] Qiu, L., and I. Benbasat, "Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems", *Journal of management information systems* 25(4), 2009, pp. 145–182.
- [46] Robert, L., "Personality in the human robot interaction literature: A review and brief critique", *Robert, LP (2018). Personality in the Human Robot Interaction Literature: A Review and Brief Critique, Proceedings of the 24th Americas Conference on Information Systems, Aug, (2018)*, 16–18.
- [47] Robert, L., R. Alahmad, C. Esterwood, S. Kim, S. You, and Q. Zhang, "A Review of Personality in Human–Robot Interactions", *Available at SSRN 3528496*, 2020.
- [48] Robert, L.P., R. Alahmad, C. Esterwood, S. Kim, S. You, and Q. Zhang, "A Review of Personality in Human Robot Interactions", *arXiv preprint arXiv:2001.11777*, 2020.
- [49] Salem, M., G. Lakatos, F. Amirabdollahian, and K. Dautenhahn, "Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust", *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, IEEE (2015), 1–8.
- [50] Scherer, K.R., *Personality markers in speech*, Cambridge University Press, 1979.
- [51] Schroeder, J., and M. Schroeder, "Trusting in Machines: How Mode of Interaction Affects Willingness to Share Personal Information with Machines", *Proceedings of the 51st Hawaii International Conference on System Sciences*, (2018).
- [52] Schuetzler, R., G. Grimes, J. Giboney, and J. Nunamaker, "The Influence of Conversational Agents on Socially Desirable Responding", *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018, pp. 283–292.
- [53] Smestad, T.L., and F. Volden, "Chatbot Personalities Matters", *Internet Science*, Springer International Publishing (2019), 170–181.
- [54] Snyder, M., "The influence of individuals on situations: Implications for understanding the links between personality and social behavior", *Journal of personality* 51(3), 1983, pp. 497–516.
- [55] Strohmann, T., D. Siemon, and S. Robra-Bissantz, "Designing Virtual In-vehicle Assistants: Design Guidelines for Creating a Convincing User Experience", *AIS Transactions on Human-Computer Interaction* 11(2), 2019, pp. 54–78.
- [56] Taylor, M., and others, "Personality Styles: Why They Matter in the Workplace", *Economic Alternatives*(1), 2020, pp. 148–163.
- [57] Thoresen, J.C., Q.C. Vuong, and A.P. Atkinson, "First impressions: Gait cues drive reliable trait judgements", *Cognition* 124(3), 2012, pp. 261–271.
- [58] Thorne, A., "The press of personality: A study of conversations between introverts and extraverts.", *Journal of Personality and Social Psychology* 53(4), 1987, pp. 718.
- [59] Turing, A.M., "Computing Machinery and Intelligence", *Mind, New Series* 59(236), 1950, pp. 433–460.
- [60] Urbach, N., F. Ahlemann, and others, "Structural equation modeling in information systems research using partial least squares", *Journal of Information technology theory and application* 11(2), 2010, pp. 5–40.
- [61] Wang, H., and M. Song, "Ckmeans. 1d.dp: optimal k-means clustering in one dimension by dynamic programming", *The R journal* 3(2), 2011, pp. 29.
- [62] Wu, H., and S.-O. Leung, "Can Likert Scales be Treated as Interval Scales?—A Simulation Study", *Journal of Social Service Research* 43(4), 2017, pp. 527–532.
- [63] Yarkoni, T., "Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers", *Journal of research in personality* 44(3), 2010, pp. 363–373.
- [64] Zumstein, D., and S. Hundertmark, "Chatbots - An Interactive Technology for Personalized Communication, Transactions and Services.", *IADIS International Journal on WWW/Internet* 15(1), 2017.