Examining User Preference for Agreeableness in Chatbots

Sarah Theres Völkel
sarah.voelkel@ifi.lmu.de
LMU Munich
Munich, Germany

Lale Kaya
l.kaya@campus.lmu.de
LMU Munich
Munich, Germany

ABSTRACT

Recent research suggests that deliberately manipulating a chatbot's personality and matching it to the user's personality can positively impact the user experience. Yet, little is known about whether this similarity attraction effect also applies to the personality dimension *agreeableness*. In a lab experiment, 30 participants interacted with three versions of an agreeable chatbot (agreeable, neutral, and disagreeable). Whilst our results corroborate a similarity attraction effect between user agreeableness and their preference for the agreeable chatbot, we did not find a reversed relationship with a disagreeable chatbot. Our findings point to a need for *moderate* instead of extreme chatbot personalities.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*.

KEYWORDS

agreeableness, chatbot, conversational agent, personalisation, personality, similarity attraction

ACM Reference Format:
Sarah Theres Völkel and Lale Kaya. 2021. Examining User Preference for Agreeableness in Chatbots. In *3rd Conference on Conversational User Interfaces (CUI '21), July 27–29, 2021, Bilbao (online), Spain.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3469595.3469633

1 INTRODUCTION

Chatbots are ubiquitous, integrated in customer service, online shopping, and information retrieval applications. These conversational agents are often considered social actors [24], with users unconsciously assigning them *personalities* [28]. Zhou et al. [41] found out that deliberately manipulating this personality perception has an impact on user trust and engagement. Research on other conversational agents (CAs), such as voice assistants [5] and robots [1], corroborates these findings.

Similar to human-human interaction, users prefer certain personality types, tending to favour chatbots which share congruent personalities with them [12, 33], coined the *similarity attraction effect* [4, 23]. For example, matching user and chatbot personality had a positive impact on user engagement [33] as well as users' self-disclosure and their willingness to accept the chatbot's advice [12].

Research on the similarity attraction effect in human-CA interaction has focused on the personality dimension *Extraversion* due to its close link to behaviour [1, 12, 23]. However, other personality traits, such as *Agreeableness*, seem particularly interesting in this context, since CAs are primarily used as helpful assistants in service applications. This is echoed by previous work on modelling speech-based CA personality, which highlighted the role of service-oriented personality dimensions [39].

It is questionable whether the preference for agreeable chatbots also follows a similarity attraction effect: Whilst agreeable users are likely to favour an agreeable chatbot, disagreeable users might not expect an uncooperative, unhelpful chatbot, given that these characteristics are usually not associated with assistants. Similarly, research on human-human similarity attraction found that employees' agreeableness only affected the attitude towards an organisation for applicants high in agreeableness, not for those with low scores [36].

In order to examine the relationship between user agreeableness and their preference for agreeableness in chatbots, we developed three different versions of an agreeable chatbot (agreeable, neutral, disagreeable). To imbue the chatbot with personality, we draw upon an abundance of work in psychology and linguistics that has examined how personality is manifested through human language [15, 21, 27, 31]. Similar to previous research [23], we leverage this relationship to equip the chatbots with different language styles. Specifically we address the following two research questions:

(1) RQ1: *Can we synthesise different levels of agreeableness in a chatbot by systematically varying its language style?*
(2) RQ2: *Is there a relationship between user agreeableness and their preference for agreeableness in a chatbot?*

2 RELATED WORK

Below we summarise work on human and conversational agent personality, personality markers in language, and adapting the agent to the user.

2.1 Human and Conversational Agent Personality

Personality describes an individual's consistent and characteristic patterns of behaviour, emotions, and cognition [20]. The *Big Five* (also called *Five-Factor model* or *OCEAN*) has emerged as the most prevalent paradigm for modelling human personality in psychology research and comprises five broad dimensions: *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, *Neuroticism*. The dimension *Agreeableness* reflects a tendency to be trustful, genuine, helpful, modest, obliging, and cooperative.

CUI '21, July 27–29, 2021, Bilbao (online), Spain

ACM ISBN 978-1-4503-8998-3/21/07…$15.00
https://doi.org/10.1145/3469595.3469633

Since users treat conversational agents as if they were people [24], HCI researchers also referred to the Big Five model for describing differences in how conversational agents express behaviour [22, 32, 35]. For example, Ruane et al. [29] contrasted users' perceptions of a chatbot high in extraversion and agreeableness with a chatbot low in these two dimensions. For speech-based conversational agents, Cafaro et al. [5] manipulated the impression of an extraverted virtual museum guide.

2.2 Personality Markers in Language

The relationship between human personality and perceptible behaviour cues has long been researched in psychology and linguistics [6, 7, 9, 11, 21, 25–27, 30, 31]. People high in agreeableness are characterised by using more positive (e.g., "love", "happy") and less negative emotion words (e.g., "ugly", "hurt"), whereas this ratio is reversed for people low in agreeableness [21, 27, 40]. A person's personality also influences their choice of words. For example, an agreeable person's language comprises more family-related (e.g., "For the sake of my family.") and inclusive words (e.g., "I felt included."), consistent with agreeable people's predisposition towards strong social relationships [15]. Moreover, agreeable people tend to use more certainty-related words (e.g., "I felt total security.") [15]. Conversely, people low in agreeableness tend to slip in more swear words (e.g., "Damn!") [16, 27] and mannerisms (e.g., "You know") [31], along with expressions of anger (e.g., "I hate school.") [15]. Furthermore, the use of emojis is associated with personality traits [19, 38]. Agreeable people were found to use more blushing [19] and heart-related emojis, such as kissing faces and hearts [38].

These insights have been leveraged to imbue chatbots with an intended personality through design. For example, Ruane et al. [29] synthesised two versions of a chatbot by imitating human choice of words associated with extraversion and agreeableness. Zhou et al. [41] used language styles, such as questioning style and expressive vs terse phrases, to infuse different characteristics in their chatbots.

For speech-based conversational agents and robots, these linguistic cues have been replenished with paraverbal and nonverbal behaviour manifestations. For example, Lee and Nass [17] employed different voice parameters, such as pitch, fundamental frequency, speech rate, and volume, to create the impression of an extraverted voice user interface whilst Andrist et al. [1] manipulated a robot's gaze behaviour to convey different expressions of extraversion.

Since the relationship between personality and behavioural cues is most pronounced for the dimension extraversion, previous work has mostly focused on imbuing CAs with different levels of extraversion (e.g., [17, 23]). Notably, Ruane et al. [29] successfully created the perception of a chatbot that is high in both extraversion and agreeableness at the same time but they did not evaluate the influence of user personality on chatbot preference. In contrast in our work, we manipulate three different versions of agreeableness in a chatbot and examine users' preference for the three levels.

2.3 Adapting Conversational Agents to the User

Users' individual preferences for particular conversational agent personalities have mostly been investigated for speech-based and embodied conversational agents [10, 23, 37]. For example, extraverted users were found to prefer a virtual real estate agent engaging in social talk [3] as well as an extraverted voice user interface on a book buying website [23]. Similarly, interacting with an introverted robot in a repetitive task increased introverted users' motivation [1].

For text-based conversational agents, Shumanov and Johnson [33] showed that matching the chatbot's language style (introverted vs extraverted) to the user's level of extraversion can have a positive impact on user engagement and increased purchases in a commerce interaction. In a similar context, matching the chatbot's and user's dominance levels resulted in perceptions of similarity, increasing users' self-disclosure of personal information during the conversation [12]. In our work, we want to examine if this similarity attraction effect can also be found for the personality dimension *agreeableness* in human-chatbot interaction.

3 DEVELOPING A PERSONALITY-IMBUED CHATBOT

To examine user preference for different levels of agreeableness, we created three chatbots, situated in a web movie recommender system. We chose this use case since chatbots are often employed in customer service and we expected users to find the situation relatable.

3.1 Conversation Flow

In order to determine the conversation flow between user and chatbot, we asked N=5 streaming service users what questions they would expect from a movie recommender chatbot. Four questions emerged from the interviews: user's preferred genre, available time, mood, and company. The dialogue is therefore structured as follows: The chatbot (1) welcomes the user, (2) asks the aforementioned questions, (3) gives a movie recommendation based on the user's preferences, (4) with the user then either accepting it or asking for another one, (5) the chatbot says goodbye.

3.2 Personality Manipulation

We deliberately manipulated the chatbot's language to create three distinctive versions, each representing a different level of agreeableness (*agreeable, neutral, disagreeable*). Similar to previous studies creating personality in robots or voice assistants [1, 5, 23], we leveraged verbal cues which are associated with *human* agreeableness to imbue the chatbot with personality. Figure 1 shows an example excerpt from the dialogue, along with its implementation in the three chatbot versions.

Since agreeable people are characterised by being friendly, empathetic, and willing to help [13], the **agreeable chatbot** agrees with the user on opinions and expresses interpersonal concern for the user. Notably, the agreeable chatbot employs positive emotions words such as "nice" or "like" [27], family-related word such as "together" or "family" [15, 40], words indicating certainty such as "I'm sure" [15], as well as blushing [19] and kiss [38] emojis, as informed by previous research.

Conversely, the **disagreeable chatbot** is pugnacious, critical, uncooperative, and does not show any interest in the user [8], e.g., not asking for the user's current mood. On top of that, it is equipped with negative emotion words such as "bad" [27], swear words such

Figure 1: Excerpts from the agreeable chatbot (left), neutral chatbot (middle), and disagreeable chatbot (right).

as "crap" [16], mannerisms such as "so" and "okay" [31], along with expressions of anger ("I'm getting angry.") [15].

Finally, the **neutral chatbot** showcases a neutral and polite language, neither expressing positive nor negative emotions in contrast to the other two versions. Moreover, it does not show a reaction to the user's choices, yet communicates in a respectful and professional way.

3.3 Implementation

We implemented the three chatbot versions on Botpress[1], an open source development platform for CAs. For the purpose of our study, we implemented a simple content-based recommender system, based on users' preferences and context. The data base consisted of 30 movies, namely the three best-rated movies for ten common genres, as informed by a German movie recommender website[2].

4 RESEARCH DESIGN

We conducted a within-groups lab experiment to investigate our research questions. In all three chatbot interactions, participants were asked to inquire from the chatbot about a movie recommendation. In each run, we slightly varied the task description to avoid the monotonous repetition of the task. This variation of the task description referred to the company, with which the user is watching the movie. Both the order of the chatbots as well as the task descriptions were counterbalanced using a Latin Square.

After each interaction, we asked participants to describe their impression of the chatbot. First, participants specified their perception of the chatbot's agreeableness by filling out the agreeableness items of the German version [8] of the Big Five Inventory-2 (BFI-2) questionnaire [34]. This established personality questionnaire [2] comprises twelve Likert scale items for each personality trait and assesses agreeableness on three facets (namely, *compassion*, *respectfulness*, and *trust*). Second, participants indicated how much they would like to interact with this chatbot again.

At the end of the study, we collected participants' self-reported level of agreeableness via the same BFI-2 personality questionnaire [8] as well as their overall ranking of the three chatbots.

[1] https://botpress.com
[2] www.moviepilot.de

We recruited participants using convenience sampling, mailings lists, and social media. On average, each session took about 30 minutes. Participants were compensated for their effort with € 5 in cash or study course credits. Our sample consisted of N=30 participants (50% male, 50% female, age range: 18–54 years, 84% of participants between 18 and 29 years old). Participants' agreeableness scores ranged from 2.08 to 4.67 with a mean score of 3.85 (SD=0.56), which is comparable to the distribution of agreeableness scores in the German population [8].

5 RESULTS

Below we present our results on the effectiveness of the agreeableness manipulation, users' desire to interact with the three chatbot versions, and a potential similarity attraction effect between user personality and preference for the chatbot personality.

5.1 Agreeableness Manipulation Check

Figure 2 shows participants' evaluation of the three chatbots' levels of agreeableness. In line with the personality questionnaire instructions [8, 34], we calculated the mean agreeableness score for each chatbot (scale from *1=disagreeable* to *5=agreeable*). Overall, the manipulation was successful, with the agreeable chatbot being perceived as more agreeable (M = 4.54, SD = 0.30) than the neutral chatbot (M = 4.06, SD = 0.63), and the disagreeable chatbot (M = 1.53, SD = 0.34). However, participants found the neutral chatbot also rather agreeable. A Greenhouse-Geisser corrected repeated-measures ANOVA underpins these results, pointing to significant differences between the three versions ($F(1.67, 48.33) = 381.12$, $p < .001$, $\eta^2 = 0.93$). Pairwise post-hoc tests yielded significant differences between all three pairs ($p < .001$).

5.2 Desire to Interact with Chatbots

As shown in Figure 3, participants preferred interacting again with the agreeable (M = 4.13, SD = 0.97) and neutral (M = 4.07, SD = 1.05) chatbots, whilst the desire to chat with the disagreeable version was rather low on average (M = 1.70, SD = 1.09). Since the data was not normally distributed, we conducted a Friedman test, which determined a significant effect of the chatbot on participants' desire to interact with the chatbot ($\chi^2(2) = 36.94$, $p < .001$). Pairwise Nemenyi post-hoc tests yielded significant differences between the preference for the agreeable and disagreeable chatbots ($p = .001$)

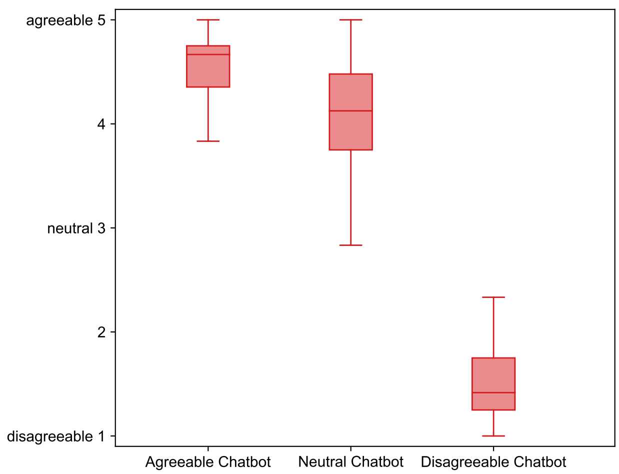

Figure 2: Participants evaluated the three chatbots regarding their perceived level of agreeableness.

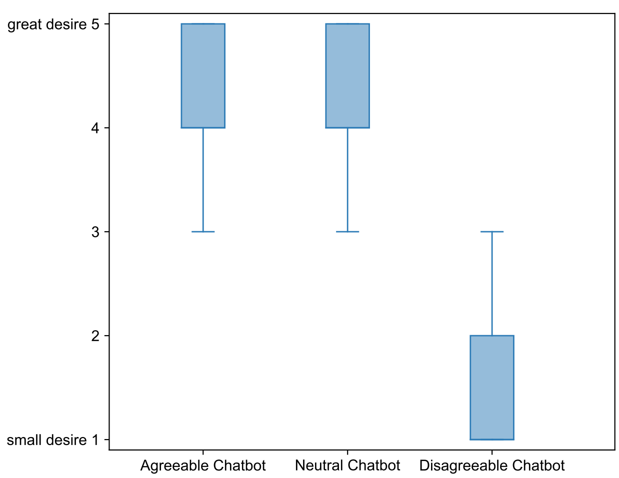

Figure 3: Participants evaluated the three chatbots regarding how much they would like to interact with the chatbot again.

as well as between the preference for the neutral and disagreeable chatbots (p = .001). There was no significant difference between participants' desire to interact with the agreeable or neutral chatbot.

These findings are supported by participants' ranking of the three chatbots at the end of the study. The agreeable chatbot was favoured by 53.3% of participants as a future interlocutor, the neutral chatbot by 43.3%, and the disagreeable chatbot by 3.3% of participants. Whilst agreeable and neutral chatbot seem to share the ranks one and two, the disagreeable chatbot was ranked worst by 86.7% of participants.

5.3 Similarity Attraction Effect

In order to evaluate whether participants preferred a chatbot with a matching level of agreeableness, we calculated Spearman's rank correlation coefficients (ρ) between participant and chatbot personality. Our results demonstrated a significant, moderate positive relationship between participants' agreeableness and their preference for the agreeable chatbot (ρ = 0.47, p = .008). The correlations between participants' agreeableness and their desire to interact with the neutral (ρ = 0.18, p = .342) and disagreeable chatbots (ρ = -0.09, p = .635) were not significant.

6 LIMITATIONS

Our method and findings are limited in several ways and should be understood with these limitations in mind.

First, whilst participants' agreeableness scores correspond to the German average agreeableness scores [8], a bigger sample representing a wider distribution of agreeableness is vital, yet difficult to achieve since volunteers' agreeableness in empirical studies are often skewed towards positive scores [18].

Second, agreeable people are prone to more acquiescent and positive response styles, probably due to social desirability [14]. Thus, agreeable participants might have evaluated all chatbots higher due to their response style in contrast to more disagreeable participants.

Third, the Big Five model was derived from human language use to describe *human* personality. Recent work by Völkel et al. [39] indicated that this model is not applicable to describe *speech-based* conversational agent personality. Instead, they presented ten alternative dimensions which comprise several subfacets of agreeableness. Whilst our findings point to a successful perception of agreeableness in chatbots, other theoretical models to specifically describe chatbot personalities might be more suitable.

7 DISCUSSION AND FUTURE WORK

Our findings show that a chatbot's level of agreeableness can be deliberately manipulated by systematically varying its language, confirming previous work by Ruane et al. [29]. However, designing a "neutral" chatbot turned out to be more difficult; participants perceived this chatbot also as rather agreeable. It is also not clear whether a *truly neutral* chatbot can be designed at all since the chatbot's task is to assist the user.

Whilst we found a similarity attraction effect between user agreeableness and preference for the agreeable chatbot, our results do not indicate a reversed relationship between user agreeableness and their preference for the disagreeable chatbot, echoing work on human-human interaction [36]. Notably, however, not all participants favoured the agreeable chatbot, either. Instead, when juxtaposing the three chatbots, participants liked the agreeable and neutral (perceived as moderately agreeable) chatbots almost equally.

A reason for this result could be that people are more used to moderate personalities in humans; the distribution of human personality also follows a normal distribution, with medium scores of a personality dimension being more frequent than the extremes [8]. However, little is known about how fine-grained levels of *moderate* agreeableness can be implemented in conversational agents since associations between "medium" levels of personality traits and their behaviour have not been the focus of psycho-linguistics until now. The reason for this lack of knowledge is that the relationship between human personality and behaviour cues is usually calculated via correlations (e.g., [15, 27]). Thus, little is known as to which thresholds define different levels of personalities. For example, does a moderately agreeable chatbot (in contrast to a highly agreeable chatbot) use positive emotion words only in every third or seventh sentence? As a consequence, previous work on user preference for personality in conversational agents has usually examined dichotomous types such as "extravert" and "introvert" [1, 5, 29], which tend to represent extremes of the respective personality dimension.

Apart from individual user preferences, future work should also examine the context of use. Whilst it is likely that in many current customer service-oriented use cases, the majority of users will expect an agreeable chatbot, there might be other scenarios in which users enjoy a more moderately agreeable agent. For example, a moderately agreeable chatbot might be more preferable to users for time-critical or repetitive tasks whereas the more verbose, highly agreeable chatbot could be favoured for chatty leisure-related tasks.

8 CONCLUSION

Recent research suggests that deliberately manipulating a conversational agent's personality and matching it to the user's personality can positively impact the user experience. Yet, little is known about whether this similarity attraction effect also applies to solely text-based conversational agents, namely chatbots, and to the personality dimension *agreeableness*. Our findings show that different levels of agreeableness can be infused in a chatbot through systematic variations in its language. Whilst we found a similarity attraction effect between user agreeableness and preference for the agreeable chatbot, our results did not indicate a reversed relationship between user agreeableness and their preference for the disagreeable chatbot. Instead, our findings point to a need for moderate personality expressions in conversational agents. Future work should therefore undertake a more fine-grained evaluation of user preference for different levels of agreeableness and examine the influence of the respective context on user preference.

REFERENCES

[1] Sean Andrist, Bilge Mutlu, and Adriana Tapus. 2015. Look Like Me: Matching Robot Personality via Gaze to Increase Motivation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3603–3612. https://doi.org/10.1145/2702123.2702592

[2] Jeromy Anglim, Sharon Horwood, Luke D Smillie, Rosario J Marrero, and Joshua K Wood. 2020. Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin* 146, 4 (2020), 279–323. https://doi.org/10.1037/bul0000226

[3] Timothy W Bickmore and Rosalind W Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)* 12, 2 (2005), 293–327.

[4] Donn Erwin Byrne. 1971. *The attraction paradigm.* Academic Press, Cambridge, MA, USA.

[5] Angelo Cafaro, Hannes Högni Vilhjálmsson, and Timothy Bickmore. 2016. First Impressions in Human–Agent Virtual Encounters. *ACM Trans. Comput.-Hum. Interact.* 23, 4, Article 24 (Aug. 2016), 40 pages. https://doi.org/10.1145/2940325

[6] Anne Campbell and J. Philippe Rushton. 1978. Bodily communication and personality. *British Journal of Social and Clinical Psychology* 17, 1 (1978), 31–36. https://doi.org/10.1111/j.2044-8260.1978.tb00893.x

[7] DW Carment, CG Miles, and VB Cervin. 1965. Persuasiveness and persuasibility as related to intelligence and extraversion. *British Journal of Social and Clinical Psychology* 4, 1 (1965), 1–7. https://doi.org/10.1111/j.2044-8260.1965.tb00433.x

[8] Daniel Danner, Beatrice Rammstedt, Matthias Bluemke, Lisa Treiber, Sabrina Berres, Christopher Soto, and Oliver John. 2016. *Die deutsche Version des Big Five Inventory 2 (BFI-2).* GESIS, Mannheim, Germany. https://doi.org/10.6102/zis247

[9] Jean-Marc Dewaele and Adrian Furnham. 2000. Personality and speech production: A pilot study of second language learners. *Personality and Individual Differences* 28, 2 (2000), 355–365. https://doi.org/10.1016/S0191-8869(99)00106-3

[10] Patrick Ehrenbrink, Seif Osman, and Sebastian Möller. 2017. Google Now is for the Extraverted, Cortana for the Introverted: Investigating the Influence of Personality on IPA Preference. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (Brisbane, Queensland, Australia) *(OZCHI '17)*. Association for Computing Machinery, New York, NY, USA, 257–265. https://doi.org/10.1145/3152771.3152799

[11] Adrian Furnham. 1990. Language and Personality. In *Handbook of language and social psychology*, William Peter Robinson and Howard Giles (Eds.). John Wiley & Sons, Chichester, UK, 73–95.

[12] Ulrich Gnewuch, Meng Yu, and Alexander Maedche. 2020. The Effect of Perceived Similarity in Dominance on Customer Self-Disclosure to Chatbots in Conversational Commerce. In *Proceedings of the 28th European Conference on Information Systems (ECIS 2020)*. AIS, eLibrary (AISeL).

[13] William G Graziano, Meara M Habashi, Brad E Sheese, and Renée M Tobin. 2007. Agreeableness, empathy, and helping: A person × situation perspective. *Journal of Personality and Social Psychology* 93, 4 (2007), 583–599. https://doi.org/10.1037/0022-3514.93.4.583

[14] Jia He and Fons J.R. van de Vijver. 2013. A general response style factor: Evidence from a multi-ethnic study in the Netherlands. *Personality and Individual Differences* 55, 7 (2013), 794–800. https://doi.org/10.1016/j.paid.2013.06.017

[15] Jacob B. Hirsh, Colin G. DeYoung, and Jordan B. Peterson. 2009. Metatraits of the Big Five Differentially Predict Engagement and Restraint of Behavior. *Journal of Personality* 77, 4 (2009), 1085–1102. https://doi.org/10.1111/j.1467-6494.2009.00575.x

[16] Thomas Holtgraves. 2011. Text messaging, personality, and the social context. *Journal of Research in Personality* 45, 1 (2011), 92 – 99. https://doi.org/10.1016/j.jrp.2010.11.015

[17] Kwan Min Lee and Clifford Nass. 2003. Designing Social Presence of Social Actors in Human Computer Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 289–296. https://doi.org/10.1145/642611.642662

[18] Jan-Erik Lönnqvist, Sampo Paunonen, Markku Verkasalo, Sointu Leikas, Annamari Tuulioâ-Henriksson, and Jouko Lönnqvist. 2007. Personality characteristics of research volunteers. *European Journal of Personality* 21, 8 (2007), 1017–1030. https://doi.org/10.1002/per.655

[19] Davide Marengo, Fabrizia Giannotta, and Michele Settanni. 2017. Assessing personality using emoji: An exploratory study. *Personality and Individual Differences* 112 (2017), 74 – 78. https://doi.org/10.1016/j.paid.2017.02.037

[20] Robert R. McCrae and Paul T. Costa. 2008. A five-factor theory of personality. In *Handbook of Personality: Theory and Research*, O.P. John, R.W. Robins, and L.A. Pervin (Eds.). Vol. 3. The Guilford Press, New York, NY, USA, 159–181.

[21] Matthias R Mehl, Samuel D Gosling, and James W Pennebaker. 2006. Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology* 90, 5 (2006), 862–877. https://doi.org/10.1037/0022-3514.90.5.862

[22] Clifford Nass and Scott Brave. 2005. *Wired for speech: How voice activates and advances the human-computer relationship.* MIT press, Cambridge, MA, USA.

[23] Clifford Nass and Kwan Min Lee. 2001. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied* 7, 3 (2001), 171. https://doi.org/10.1037/1076-898X.7.3.171

[24] 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* (Boston, Massachusetts, USA) *(CHI '94)*. Association for Computing Machinery, New York, NY, USA, 72–78. https://doi.org/10.1145/191666.191703

[25] Jon Oberlander and Alastair J. Gill. 2004. Individual differences and implicit language: personality, parts-of-speech and pervasiveness. *Proceedings of the Annual Meeting of the Cognitive Science Society* 26 (2004), 1035–1040.

[26] M. Patterson and D.S. Holmes. 1966. Social Interaction Correlates of the MMPI Extraversion Introversion Scale. *American Psychologist* 21 (1966), 724–25.

[27] James W. Pennebaker and Laura A. King. 1999. Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology* 77, 6 (1999), 1296–1312. https://doi.org/10.1037/0022-3514.77.6.1296

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

[29] Elayne Ruane, Sinead Farrell, and Anthony Ventresque. 2021. User Perception of Text-Based Chatbot Personality. In *Chatbot Research and Design*. Springer International Publishing, Cham, Switzerland, 32–47. https://doi.org/10.1007/978-3-030-68288-0_3

[30] D. R. Rutter, Ian E. Morley, and Jane C. Graham. 1972. Visual interaction in a group of introverts and extraverts. *European Journal of Social Psychology* 2, 4 (1972), 371–384. https://doi.org/10.1002/ejsp.2420020403

[31] Klaus Rainer Scherer. 1979. Personality markers in speech. In *Social Markers in Speech*, Klaur Rainer Scherer and Howard Giles (Eds.). Cambridge University Press, Cambridge, UK.

[32] Michael Schmitz, Antonio Krüger, and Sarah Schmidt. 2007. Modelling Personality in Voices of Talking Products through Prosodic Parameters. In *Proceedings of the 12th International Conference on Intelligent User Interfaces* (Honolulu, Hawaii, USA) *(IUI '07)*. Association for Computing Machinery, New York, NY, USA, 313–316. https://doi.org/10.1145/1216295.1216355

[33] Michael Shumanov and Lester Johnson. 2021. Making conversations with chatbots more personalized. *Computers in Human Behavior* 117 (2021), 106627. https://doi.org/10.1016/j.chb.2020.106627

[34] Christopher J. Soto and Oliver P. John. 2017. The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology* 113, 1 (2017), 117 – 143. https://doi.org/10.1037/pspp0000096

[35] Jürgen Trouvain, Sarah Schmidt, Marc Schröder, Michael Schmitz, and William J. Barry. 2006. Modelling personality features by changing prosody in synthetic speech. In *Proceedings of the 3rd International Conference on Speech Prosody*. TUDpress, Dresden, Germany, 4 pages. https://doi.org/10.22028/D291-25920

[36] Greet Van Hoye and Daniel B. Turban. 2015. Applicant-Employee Fit in Personality: Testing predictions from similarity-attraction theory and trait activation theory. *International Journal of Selection and Assessment* 23, 3 (2015), 210–223. https://doi.org/10.1111/ijsa.12109

[37] Sarah Theres Völkel, Daniel Buschek, Malin Eiband, Benjamin R. Cowan, and Heinrich Hussmann. 2021. Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 254, 15 pages. https://doi.org/10.1145/3411764.3445536

[38] Sarah Theres Völkel, Daniel Buschek, Jelena Pranjic, and Heinrich Hussmann. 2019. Understanding Emoji Interpretation through User Personality and Message Context. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) *(MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article 3, 12 pages. https://doi.org/10.1145/3338286.3340114

[39] Sarah Theres Völkel, Ramona Schödel, Daniel Buschek, Clemens Stachl, Verena Winterhalter, Markus Bühner, and Heinrich Hussmann. 2020. Developing a Personality Model for Speech-Based Conversational Agents Using the Psycholexical Approach. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376210

[40] Tal Yarkoni. 2010. Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality* 44, 3 (2010), 363 – 373. https://doi.org/10.1016/j.jrp.2010.04.001

[41] Michelle X. Zhou, Gloria Mark, Jingyi Li, and Huahai Yang. 2019. Trusting Virtual Agents: The Effect of Personality. *ACM Trans. Interact. Intell. Syst.* 9, 2-3, Article 10 (March 2019), 36 pages. https://doi.org/10.1145/3232077