

We are all Individuals: The Role of Robot Personality and Human Traits in Trustworthy Interaction

Mei Yii Lim¹, José David Aguas Lopes², David A. Robb¹, Bruce W. Wilson¹,
Meriam Moujahid¹, Emanuele De Pellegrin¹ and Helen Hastie¹

Abstract—As robots take on roles in our society, it is important that their appearance, behaviour and personality are appropriate for the job they are given and are perceived favourably by the people with whom they interact. Here, we provide an extensive quantitative and qualitative study exploring robot personality but, importantly, with respect to individual human traits. Firstly, we show that we can accurately portray personality in a social robot, in terms of extroversion-introversion using vocal cues and linguistic features. Secondly, through garnering preferences and trust ratings for these different robot personalities, we establish that, for a Robo-Barista, an extrovert robot is preferred and trusted more than an introvert robot, regardless of the subject's own personality. Thirdly, we find that individual attitudes and predispositions towards robots do impact trust in the Robo-Baristas, and are therefore important considerations in addition to robot personality, roles and interaction context when designing any human-robot interaction study.

I. INTRODUCTION

As robots become more capable of meaningful interaction, we will start to see them in settings that require social etiquette, such as in the service industry, including as baristas [1], receptionists [2] and bartenders [3]. As these are everyday settings, they provide convenient platforms for experiments that do not require subjects to have specialist training, such as in the first responder or healthcare domains. Robots that serve in more of a social setting, such as service robots, as well as those used more as tools, such as surveying drones, all need to be trusted to do their job well [4]. The Robo-Barista scenario discussed here, thus facilitates long term “in the wild” Human Robot Interaction (HRI) experimentation, allowing for the exploration of human-robot trust over time. Subsequently, this will, hopefully, encourage the adoption and acceptance of robots in our everyday lives.

Prior work has shown that both robot and human personality are significant predictors of trust [5]. According to Jooisse et al. [6], people implicitly assume that certain roles or occupations require certain personalities, but which personality would be appropriate for our Robo-Barista? When Howarth [7] asked subjects to list occupations under extrovert and introvert categories, he found that service industry jobs such as waiter/ess, steward/ess and bartender were all listed under the extrovert categories. Furthermore, Sandstrom and Dunn

[8] found that people who interacted with a human barista as if they were a “weak tie-friend”, instead of a stranger, experienced more positive affect, mediated by a feeling of belonging. This led us to explore the research question of whether this finding maps over to HRI. Specifically, we explore here if a more extroverted Robo-Barista will be preferred and trusted over an introvert one.

Trust in HRI is related to trust in automation and varying levels of trust in robots exist, which might result in misuse (over reliance in cases of extremely high levels of trust) or disuse of systems and robots (in cases of very low levels of trust) [9]. Furthermore, unfounded fluctuations in trust could inhibit user acceptance, thereby compromising the inherent advantages of technology [10]. It is also important to take into consideration the person's attitude and propensity to trust, as well as the context of interaction and the task at hand [5], [11].

As well as the commonly used ‘Negative Attitude to Robots’ (NARS) measure [12], we also explore the subject's Propensity to trust (PTT), which is a relatively stable disposition that is developed over time. It can lead to trusting beliefs, i.e. perceived trustworthiness, which are formed as a trustor observes, interprets and ascribes motives to a trustee's actions [13], especially in novel situations [14]. Rotter [15] studied PTT in human-human interaction, while Jessup et al. [16] explored PTT in human-automation interaction. Lewis et al. [17] highlighted that there is very little empirical work that has investigated how users' PTT affects their relationship with a robot. In order to design robots for a specific context, continuing research on how, and to what extent, PTT plays a role in human-robot trust is necessary, hence it is one of the focuses of this paper.

As well as describing our general experimental platform for a Robo-Barista (Fig. 1), we address the following research questions:

- **R1:** Can extroversion and introversion be obviously implemented as personalities for a social robot, through linguistic and prosodic cues?
- **R2:** Which personality (extrovert or introvert) is preferred and trusted more for a service robot, such as a Robo-Barista? Do similarity/complementary-attraction theories between users and robots apply to a Robo-Barista?
- **R3:** Are preference and trust measures correlated?
- **R4:** Is trust in a service robot sensitive to individual differences and if so which ones?

This work was funded and supported by the UKRI Node on Trust (EP/V026682/1).

¹School of Mathematical and Computer Sciences, Heriot-Watt University, EH14 4AS, UK m.lim@hw.ac.uk, d.a.robb@hw.ac.uk, bww1@hw.ac.uk, mm470@hw.ac.uk, ed50@hw.ac.uk, h.hastie@hw.ac.uk

²Semasio, Portugal jose@semasio.com

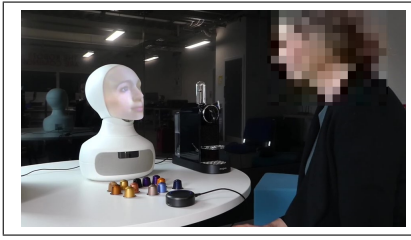


Fig. 1. The Robo-Barista interacting with a user

II. BACKGROUND

While there is no generally agreed-upon definition of personality, it refers to traits that predict a person’s behaviour and one of the most used instruments is the Big Five theory, which proposes 5 personality types: extroversion, agreeableness, conscientiousness, neuroticism and openness to experience [18]. It has been found to be robust and has external validity. Extraversion has been well-established to be positively related to dispositional interpersonal trust [11], [19] and was found to generalise to PTT in machines [20], [21]. Much research on robot personality has been focused on the dimension of extroversion, as it is one of the easier traits to display, may be more salient in shorter interactions [22] and is, therefore, why we focused the robot’s personality dimension along this axis.

Modulating prosody is an effective way to portray the extroversion-introversion dimensions in artificial agents [23], [24] and has a strong influence on the perceived personality. In general, extroverts speak with louder voice, higher fundamental frequency, broader frequency range and faster speech rate than introverts [25], [26]. It is these prosodic parameters - volume, pitch and tempo - that we manipulate to create our two robot personalities of extrovert and introvert. Furthermore, there are pragmatic and lexical/syntactic differences [27], [28]. For example introverts use broader vocabulary, while extroverts are more repetitive and have fewer pauses and hesitations [29].

Lohse et al. [30] found significant preference for extrovert robot behaviour, which was rated more friendly, diversified, intelligent and faster. With regards perceived intelligence, Leuwerink’s [31] study showed that introvert robots are considered more intelligent in group interactions, while extrovert robots are more intelligent in dyadic interaction. Importantly, Alarcon et al. [32] demonstrate that individual differences play a role in the trust process. Studies have lent support for the idea that humans lean towards robots that have similar traits to them, an effect known as similarity-attraction [33]. However, this effect has not been unanimously demonstrated, with some studies showing there may be more of a complementary-attraction effect, that is, a preference for agents that complement their personality [34] but do not necessarily match it. Joosse et al.’s work [6] does align more with the complementary-attraction effect, but they posit that it depends on the interaction context and the robot role. This adds further complexity to the personality matching debate. Rousseau et al. [35] also stress the importance of context

to understanding trust. Currently, there are not many studies on anthropomorphic robots in this area [36] and none in the context of a robot barista that we are aware of. This has led us to explore **R2** and **R3**.

With regards to trustworthy characteristics, Hancock et al. [5] in their meta-review, found that robot-related, human-related and contextual factors are predictors of trust. Additionally, language indicators such as positive, gratitude and collaborative cues have been shown to help in identifying trustworthy interaction [37]. Embodiment, social intelligence capabilities and non-verbal communication are other factors affecting trust [38]. To answer **R4**, we will focus on human-related factors, specifically whether the subjects’ Big 5 personality traits, PTT and NARS relate to perceived trustworthiness of the Robo-Barista.

III. ROBO-BARISTA SET-UP

We created a Robo-Barista to enable visitors and staff at our institution to talk and interact with a social robot, as illustrated in Fig. 1. The robot, takes drinks orders and interfaces with a coffee machine, which will automatically complete the order. The robot is able to detect when a potential customer is in the vicinity, either by the person directly addressing it or by it having detected their social presence. While it is preparing the drink, it engages the user in small talk and gives information on the institution. Finally, it alerts the user once the drink is ready and politely ends the conversation.

The Robo-Barista set-up consists of a Furhat robot from Furhat Robotics¹ connected to a Nespresso CitiZ Espresso coffee machine. The dialogue was implemented using RASA 2.0². For Natural Language Understanding, we used their DIET model [39] trained using data generated from template grammars. The DIET model handles both user intent classification and entity recognition in a single transformer architecture and allows for the use of pre-trained embedding models, such as BERT [40]. For the dialogue, we used the RASA TED Policy [41], combined with a set of handwritten dialogue rules (e.g. for greetings and farewells). The TED Policy applies a machine learning policy that can generalise patterns from example conversations using context from previous utterances in the dialogue. The user may interject with off-topic remarks and have multiple embedded sub-dialogues (e.g., when the user changes their coffee order or engages in small talk with the robot).

As mentioned above, to create the two robot personalities, the tempo, pitch and volume of the robot’s speech were manipulated based on the speech traits literature for extroversion and introversion [23]–[26]. Taking Furhat’s standard setting as a baseline, the extrovert robot speech was given a higher pitch (+20%), volume (+6dB) and speaking rate (+20%), while the introvert robot speech has a lower pitch (-20%), volume (-6dB) and speaking rate (-20%) than the baseline. The interaction content was manipulated to reflect extrovert

¹<https://furhatrobotics.com>

²<https://rasa.com>

and introvert linguistic patterns based on the parameters provided in [27], [28]. Table I lists some example statements from the extrovert and introvert robots.

TABLE I
EXTROVERT AND INTROVERT STATEMENTS AND LINGUISTIC FEATURES
APPLIED (IN SQUARE BRACKETS)

Extrovert Robo-Barista	Introvert Robo-Barista
There are many capsules available. Why not let me help you? [self-reference]	There are many capsules available. Perhaps you might need some help? [downtoner hedges]
How about I make you a Ristretto? [informal]	I can make you a Kazaar if you like? [downtoner hedges]
It would be amazing if you can place a black capsule in the machine? [emphasizer hedges]	Can you put one of the dark blue capsules in the machine?
Great! Keep exploring! There are many other robots to meet!	Make sure you do! There are many interesting and useful robots to meet! [claim complexity]
Your order is ready! Enjoy your deliciously smelling Ristretto! Bye! [subject implicitness]	Your order is ready! Enjoy your Kazaar! Goodbye.

IV. EXPERIMENT

A. Pilot

1) *Procedure*: As a manipulation check and to address **R1**, we replicated prior work showing that subjects can recognise personality from speech traits and linguistic manipulation [23], [24]. The pilot study employed a within-subject design. The procedure was ethically approved by Heriot-Watt University Department of Computer Science’s ethics board, as is the main experiment. A total of 50 subjects took part and data from 5 subjects were excluded from analysis due to too fast and thus unrealistic and implausible completion times. Subjects were aged 18 to 62 (mean = 32.11, SD = 10.81), consisting of 34 females and 11 males. Recruitment was through the online Prolific³ platform. Pre-screening filters were applied to ensure that our subjects fall in the right age range (18-65) and are not in vulnerable user groups. The experiment was administered using Qualtrics⁴.

The subjects first provided digital consent followed by a confirmation that they have audio and video capabilities on their device. Next, they provided their gender and age information. Then they were asked to complete the agent personality perception experiment, by watching 2 videos (about 90 seconds each)⁵ of a user interacting with the Robo-Barista speaking with differing speech prosody traits and linguistic features, as described in Section III. The order of the videos was randomised to counterbalance any order effect.

After each video, subjects were asked to rate 14 adjective pairs on a 5-point Likert scale, namely Semantic

Differential items [42]. The adjectives describe the Robo-Barista’s behaviours as follows: active-passive; interested-indifferent; talkative-quiet, intelligent-stupid; predictable-unpredictable; consistent-inconsistent, fast-slow; polite-impolite; friendly-unfriendly; obedient-disobedient; diversified-boring; attentive-inattentive. They also include those describing general usefulness: useful-useless; practical-impractical [30]. As a validation check, subjects were asked to rate the robot on the introvert-extrovert dimensions on a 5-point Likert scale with 1 being introvert and 5 being extrovert. The final debrief outlined that the purpose of the survey was to assess how changes in speech prosody and linguistic features modify the perception of the robot personality on the extroversion-introversion dimensions.

2) *Results*: Wilcoxon Signed-Rank tests were performed to assess differences across conditions. The following 7 adjective pairs showed significant differences between ratings for the two robot personalities: passive-active ($Z = -2.798, p = 0.05$), indifferent-interested ($Z = -0.256, p = 0.01$), quiet-talkative ($Z = -4.497, p < 0.001$), slow-fast ($Z = -5.065, p < 0.001$), unfriendly-friendly ($Z = -2.609, p = 0.09$), boring-diversified ($Z = -3.21, p = 0.001$) and inattentive-attentive ($Z = -3.3, p < 0.001$).

Importantly, a significant difference was also found in the introvert-extrovert dimension rating, $Z = -4.284, p < 0.001$. Therefore, the differences between the ratings of the introvert and extrovert robot were clearly perceived. This provides strong evidence that our manipulation of speech and linguistic features was successful.

B. Main Experiment

1) *Procedure*: The aim of the main experiment is to address the research questions raised in Section I. 202 new subjects were recruited on Prolific. They were aged 18 to 60 (mean=28.11, SD=8.55) consisting of 77 female, 122 male and 3 non-binary. Participant recruitment and pre-screening followed the same procedure as the pilot study.

Prior to the experiment, subjects signed a consent form and were asked to provide demographic information including age, gender and prior experiences and attitudes to robots using the 14-item NARS [12] and the 6-item PTT questionnaires [16]. NARS measures negative attitudes towards robots, while PTT was designed to measure stable characteristics in individuals, attitudes towards technology and whether people were likely to collaborate with technology. Both questionnaires use a 5-point Likert scale ranging from strongly disagree to strongly agree. Jessup et al. [16] have shown that adapting the PTT questionnaire to the interaction context enhances the reliability of the measure, and hence the predictability of perceived trustworthiness. Following this prior-work, the PTT questionnaire was adapted to our scenario with the word “technology” being replaced by “robot”. We added an additional question to the NARS specific to our scenario - “I would feel uneasy if I had to order a drink from robots”.

Then Big 5 44-item inventory [18] was administered. Scores for all 5 traits were calculated for each subject and

³<https://www.prolific.co/>

⁴<https://www.qualtrics.com/uk/>

⁵Introvert Robo-Barista: <https://youtu.be/BTk3d6dkIts>

⁶Extrovert Robo-Barista: <https://youtu.be/py3h8uLvq5g>

used for correlation analysis. Additionally, categorisation was performed for the extroversion trait to allow personality matching between users' and the Robo-Barista's personality. The mean extroversion score across the entire sample is $M = 23.50$. The recommended approach is to compare the subjects' score against a published norm or construct a local norm [43]. Given that there is no official published norm, we establish a local norm for our sample population. Assuming subjects scoring above mean as extroverts, while subjects scoring below mean as introverts, we have 98 extroverts and 103 introverts. Two subjects (score 13 and 20) have a missing value in one question each, so the one who scored 20 has been excluded from analysis, as they may fall in either dimension.

During the experiment, subjects watched the same two short videos as in the pilot study, presented in alternating order. After each video, we repeated the pilot manipulation check and asked them to rate the videos using a 5-point Likert scale spanning the introvert (1) and extrovert (5) dimension. Then, they were asked to provide ratings of the level of their trust in the Robo-Barista using an adapted LETRAS-G [11], a subset of the empirically determined scale of trust in automated systems [44]. The scale contains 7 items (listed below) on a 7-point Likert ranging from not at all (1) to extremely (7), designed to measure the subject's trust in a specific robot. Thus two trust scores, one for each version of the Robo-Barista were computed for analysis.

- **LG1:** I trust the robot
- **LG2:** The robot is reliable
- **LG3:** The robot is trustworthy
- **LG4:** I am sceptical about the robot - Reverse
- **LG5:** I distrust the robot's suggestions – Reverse
- **LG6:** I am suspicious of the robot - Reverse
- **LG7:** Using the robot will lead to dangerous and harmful consequences - Reverse

The subjects were also asked to rate how much they like the robot's voice ranging from strong dislike (1) to strong like (5) and whether they can hear the speech clearly ranging from not at all (1) to all the time (5). After watching and rating both videos, the subjects were asked to select: 1) their preferred robot and 2) the robot they trusted more. They were asked to explain their choices for preference and trust through worded qualitative feedback.

2) *Independent and Dependent Variables:* This study has a within-subjects design with the independent variables as below. **IV1** and **IV4** are manipulated while **IV2**, **IV3** and **IV5** are non-manipulated variables.

- **IV1:** The Robo-Barista personality (Introvert/Extrovert)
- **IV2:** The subjects' PTT score
- **IV3:** The subjects' NARS score (with the extra question specific to our scenario)
- **IV4:** The subjects' Introvert/Extrovert category
- **IV5:** The subjects' scores for each Big 5 trait (Big 5)

The dependent variables are:

- **DV1:** The subjects' perception of the Robo-Barista's personality

- **DV2:** The choice of Robo-Barista personality that the subjects preferred for interaction
- **DV3:** The choice of Robo-Barista personality that the subjects trusted more
- **DV4:** The subjects' levels of trust in the Extrovert and Introvert Robo-Baristas, as measured using the LETRAS-G instrument

V. QUANTITATIVE RESULTS

A. **R1:** Perceived Robo-Barista Personality

With regards **R1**, a Wilcoxon Signed-Rank test indicated that the Extrovert Robot was rated significantly higher on the introversion-extroversion scale than the Introvert Robot ($Z = -8.737, p < 0.001$), confirming that the speech traits and linguistic features manipulation manifested in perceivable robot personalities.

B. **R2:** Robo-Barista Personality and Preference/Trust

To answer **R2**, we performed binomial tests on subjects' choice of preferred and trusted Robo-Barista.

1) *Personality Preference:* The binomial test on preference shows that an Extrovert Robot is preferred over an Introvert Robot, $p < 0.001$ ($n=130$ for Extrovert Robot and $n=72$ for Introvert Robot).

2) *Personality Trust:* A further binomial test indicates that an Extrovert Robot is trusted more than an Introvert Robot, $p < 0.001$ ($n=125$ for Extrovert Robot and $n=77$ for Introvert Robot). Paired t-tests showed significant differences in 5 LETRAS-G items and the overall trust levels (**DV4**) between the two Robo-Baristas in favour of the Extrovert Robot, as listed in Table II.

TABLE II
PAIRED SAMPLES T-TEST FOR LIKERT SCALE RATINGS OF THE TWO ROBO-BARISTAS' ASSIGNED TRUST LEVELS USING LETRAS-G ITEMS.
* INDICATES SIGNIFICANT.

Item	Mean E	Mean I	t	p-value
LG1: Trust	4.84	4.64	2.495	0.013*
LG2: Reliability	5.22	4.88	4.257	0.000*
LG3: Trustworthy	5.00	4.74	3.018	0.003*
LG4: Scepticism	3.64	3.94	-3.030	0.003*
LG5: Distrust	2.98	3.08	-1.273	0.205
LG6: Suspicion	3.10	3.28	-2.123	0.035*
LG7: Dangerous	2.50	2.45	0.642	0.522
Overall	4.98	4.79	3.804	0.000*

3) *Similarity/Complementary-Attraction:* Fig. 2 shows preference and trust results grouped by the subject's Introvert/Extrovert category. A Chi-square test for independence showed a non-significant association between the preferred Robo-Barista personality (**DV2**) and the subjects' Introvert/Extrovert category (**IV4**), $\chi^2(4) = 1.14, p = 0.286$. Subjects' Introvert/Extrovert category also showed no influence on the choice of more trusted Robo-Barista personality, confirmed by a Chi-square test, $\chi^2(4) = 0.358, p = 0.55$. These results do not, therefore, show evidence for the similarity/complementary-attraction theories, answering **R2**.

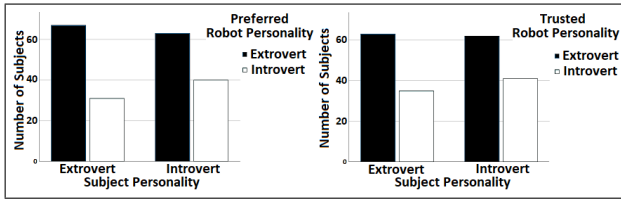


Fig. 2. “Preferred” - Left chart (DV2) and “More Trusted” - Right chart (DV3) Robo-Barista personality based on subjects’ binary choice answers grouped by subjects’ Introvert/Extrovert category (IV4)

C. R3: Preference and Trust Correlation

To address R3, a Chi-square test for independence was computed to determine whether the choice of Robo-Barista personality that the subjects trust more (DV3) relates to the choice of Robo-Barista personality the subjects preferred (DV2). An association was observed, $\chi^2(4) = 109.24, p < 0.001$. A Pearson correlation established that the choice of preferred Robo-Barista correlates positively with the choice of more trusted Robo-Barista $r(202) = 0.735, p < 0.001$.

D. R4: Trust and Individual Differences

This section reports results looking at a subject’s Big 5 traits and characteristics, such as PTT and NARS, and exploring if these affect the way they rate the trust levels of the two variations of the robot.

1) *Trust Scores and Big 5 traits*: We explore the link between subjects’ Big 5 (IV5) and their trust scores (DV4), separated out into how they rated the Extrovert and Introvert Robot. Subjects with missing values for a trait were excluded from analysis. Whilst there was no clear signal for most of the traits, we did find some interesting results. Specifically, a Pearson correlation coefficient showed positive correlation between Extrovert Robot trust score and subject agreeableness $r(199) = 0.211, p < 0.05$. Moreover, a near significant correlation was observed between Introvert Robot trust score and subject neuroticism $r(201) = -0.135, p = 0.056$.

2) *Trust Scores and PTT*: Next, we explore the link between subjects’ PTT (IV2) and their trust scores (DV4). Subjects’ PTT and trust score for the Extrovert Robot were found to be significantly positively correlated, $r(202) = 0.478, p < 0.001$. A significant positive correlation was also found between subjects’ PTT and trust score for the Introvert Robot, $r(202) = 0.484, p < 0.001$.

3) *Trust Scores and NARS*: We also investigated if subjects’ general negative attitudes to robots (IV3) is somehow linked to their trust scores (DV4). A Pearson correlation coefficient was computed for NARS and trust level in the Extrovert Robot, $r(202) = -0.554, p < 0.001$; as well as NARS and trust level in the Introvert Robot, $r(202) = -0.584, p < 0.001$ resulting in significant negative correlation in both cases.

E. Voice and Speech of Robo-Baristas

To investigate subjects’ opinion on the Robo-Baristas’ voice and speech, Wilcoxon Signed-Rank tests were conducted to compare their 5-point Likert scale ratings on how

much they like the Introvert and Extrovert Robots’ voice and how clearly they can hear the speech. The results indicated that subjects liked the Extrovert Robot voice more than the Introvert Robot voice ($Z = -4.754, p < 0.001$), while the Introvert Robot speech was heard more clearly than the Extrovert Robot speech ($Z = -4.286, p < 0.001$).

F. Confounding Variables

Mixed ANOVAs were conducted to explore whether there was an ordering effect between: ExtrovertIntrovert (those who watched the extrovert Robo-Barista video first) and IntrovertExtrovert (those who watched the introvert Robo-Barista video first) and the following variable: Robo-Barista personality ratings (Introvert, Extrovert); gender (male, female, non-binary) and Robo-Barista personality ratings; as well as age (10-19, 20-29, 30-39, 40-49, 50+) and Robo-Barista personality ratings. There was a statistically significant difference between ordering groups and Robo-Barista personality ratings, $F(1, 200) = 10.631, p = 0.001$. A Bonferroni corrected post-hoc test revealed that the IntrovertExtrovert group ratings of the Robo-Barista personality were statistically significantly higher than the ratings of the ExtrovertIntrovert group, $p = 0.001$. One possible explanation might be due to them rating the Introvert Robot quite highly on the introvert-extrovert dimension at the start resulting in them giving an even higher rating for the Extrovert Robot later. No significant interaction effect was found between order groups and personality ratings. For gender and age effect, no significant main effect was found, $F(2, 199) = 2.072, p = 0.129$ and $F(4, 197) = 0.816, p = 0.516$ respectively. The interaction effects between order groups, gender, age and personality ratings were also non-significant, ruling them out as confounding variables in our study.

VI. QUALITATIVE RESULTS

A. Codebook

Qualitative analysis was carried out on the reasons subjects provided for their robot preference and for trusting more in a specific Robo-Barista. A codebook was developed by two senior researchers, using an inductive approach similar to that used in grounded theory [45]. The initial codebook was then tested by the two researchers independently coding 10% of the questionnaires and disagreements were discussed. A team of 2 coders was then established and a small sample of questionnaires were coded individually, disagreements were resolved and clarifications were added to the codebook. Finally, a random sample of 30% of the questionnaires were dual coded independently. The average agreement across all codes was 99% with a Cohen’s Kappa of 0.93.

The codebook contains codes referring to the Robo-Baristas’ attributes that gave rise to subjects’ preference and trust (full list in Table III). Further codes pertain to the sense and sentiment of these attributes (i.e. if the subject saw an attribute as good or bad, whether the focus was towards the preferred robot or away from that not preferred, and the valence of the attribute e.g. pleasant/unpleasant). Lastly,

indifference expression where subjects were impartial about the Robo-Baristas was also coded.

TABLE III
PERCENTAGE OF ALL SUBJECTS INDICATING CERTAIN ATTRIBUTES AS THEIR PREFERENCE AND TRUST REASONS

Attribute	Preference	Trust
Appeal	22.3	12.4
Better	6.9	4
Competence	3	13.9
Ease of Communication	9.9	11.4
Holistic View of Demeanour	4	6.9
Mood	29.2	28.2
Subjects Affected or Moved	3	9.9
Tempo of Speech	11.4	5.4
Trustworthiness	1.5	8
Understandability	21.3	11.9
Voice	77.7	60.4

B. Indifference Expression

It can be observed that all subjects (n=202) expressed a reason for preference with Voice being the main determinant (77.7%). When it comes to trust, 8.4% expressed no inclination towards either Robo-Barista stating that they trusted both equally. Voice again was the main factor for trust (60.4%). Voice was the main manipulated attribute differentiating both Robo-Baristas, so this finding was as expected.

C. Attributes Reference

Next, we looked at Holistic View of Demeanour of the Robo-Baristas, indicating that the subject was referring to something other than only the voice. 4% of all subjects indicated this as their preference reason while 6.9% indicated this as their trust reason. e.g. *“It seemed more thoughtful”* [P185] referring to the Introvert Robot, which speaks slower and with lower pitch, *“Looks like a normal person being polite”* [P91] referring to the Extrovert Robot.

Looking at Appeal (enjoyable, likeability, pleasantness/niceness), Better (something unspecified about the robot is better) and Understandability of the conversations, more subjects indicated these as their preference reason than trust reason. This is mainly due to the Robo-Baristas’ voice and speech, e.g. *“It speaks in a more natural cadence...more enjoyable”* [P82], *“Clear speech, pleasant voice”* [P155].

In terms of Ease of communication, (i.e. how comfortable the subjects are at communicating with the Robo-Baristas) and Mood, (i.e. energy, friendliness, kindness, relaxed, warmth of the Robo-Baristas), the mentions for preference reason and trust reason are balanced.

D. Trust-Related Attributes

Competence of the Robo-Barista was deemed more important for trust than preference (3% for preference and 13.9% for trust reason). The perception of the Robo-Barista credibility seemed to be at play here, e.g. *“Convincing”* [P20], *“...sounds more confident...”* [P43].

This is directly linked to Trustworthiness, which was also more crucial as a trust reason than a preference reason. 8%

of subjects addressed directly, aspects of Trustworthiness, such as the Robo-Barista being more genuine, reliable or explicitly trustworthy in their trust reason but only 1.5% mention this in the preference reason. The two subthemes here were: a) those inspired to trust more the Introvert Robot due to its calm delivery and seeming reliability e.g. *“By speaking a little slower, it creates an atmosphere of trust”* [P149]; and b) those trusting more the Extrovert Robot, due to its cheerful clarity perhaps reflecting what they would expect from a Barista e.g. *“... more reliable and more normal to the standards.”* [P102], *“...upbeat attitude...”*[P46]. This perception of slow speech tempo as reliable, and fast as more normal for a Barista and thus more trustworthy, is also reflected when subjects commented on the Tempo of Speech. 5% preferred and 3.5% trusted more the slower talking Introvert Robot, while 3% preferred and 2% trusted more the faster speaking Extrovert Robot.

Another focal point for trust is how the subjects were Affected or Moved by at least one of the Robo-Baristas. 3% indicated this in their preference reason e.g. *“...it brings me peace of mind”* [P149] and *“... gives some creepy feelings”* [P24]. In contrast, 10% indicated this in their trust reason e.g. *“...made me feel more confident...”* [P37] and *“Made me feel more relaxed”* [P194]. Thus, we can see that more subjects described feeling affected in relation to their trust reason than their preference reason. This may indicate that, for many, trust was an affective feeling as opposed to a cognitive conclusion, reflecting the theoretical categorisation of trust into cognitive and affective trust [46]. Cognitive trust (someone’s readiness to rely on a service due to competence and reliability) is based on knowledge while affective trust (someone’s confidence in a service provider based on the care and concern they demonstrate) is emotional in nature.

VII. DISCUSSION

From the results, we can conclude that subjects were able to distinguish the two different Robo-Barista personalities, realised through voice and lexicon manipulation. Subjects were found to prefer an Extrovert Robot over an Introvert Robot and they also trusted the Extrovert Robot more than the Introvert Robot. This results in a positive relationship between subjects’ preferred Robo-Barista personality and the Robo-Barista personality they trusted more. The results validate the conclusion deduced from [7], [8] that extrovert Baristas will be preferred due to prior stereotype and they are deemed to be more likely to engage in small talk enabling customers to treat them like a “weak social tie” [8].

Furthermore, all LETRAS-G trust items except for two showed significant differences favouring the Extrovert Robot, consistent with the finding above. No significant differences were found for **LG5:Distrust** and **LG7:Dangerous**, which is perhaps not surprising as both robots performed their task successfully, giving no reason for subjects to distrust them and they did not portray any harmful behaviour. This might also explain why some subjects expressed indifference in their subjective responses as to which Robo-Barista they trusted more. The insignificant results of **LG5:Distrust** and

LG7:Dangerous suggest that these items are not necessarily important to our scenario, concurring with [47] that not all items in existing trust scales are always applicable but instead are context-dependent.

Interestingly, subjects' own Introvert/Extrovert personality category did not show any influence on the Robo-Barista personality they preferred or trusted more. Both introverts and extroverts preferred and trusted the Extrovert Robot more, as illustrated in Fig. 2. Although this result contradicts the personality matching theories [33], [34], it is inline with Jooisse et al.'s [6] finding that task constraints add complexity to human-robot trust relationship. One explanation could be that getting coffee is a social activity, hence the outgoing attributes of an extrovert are more favourable in general, as it conforms to existing stereotype [7].

There was no relation observed between people's predisposition to trust a robot and their own personality (Extrovert/Introvert), contradicting [20], [21] but confirming Alarcon et al.'s [32] findings. However, we did find an effect that people who are more conscientious have lower PTT, $r(199) = -0.158, p < 0.05$, thus are generally less likely to trust. In addition, people who are more agreeable showed higher perceived trustworthiness especially in the Extrovert Robot, which they trusted more. There was also a trend that people who are more neurotic had lower perceived trustworthiness in the Introvert Robot, which they trusted less. These relationships between Big 5 traits and PTT, as well as, perceived trustworthiness conform to prior work in psychology and behavioural economics [21], [48]–[50].

Perhaps unsurprising, we found that subjects' PTT in robots positively predicts their trust in the Robo-Barista. This is the case for both the Introvert and Extrovert Robots. Additionally, their prior NARS negatively predict their trust in the Robo-Baristas. As a result, subjects' PTT in robots is also negatively related to their NARS, $r(202) = -0.575, p < 0.001$. This indicates that generally those with higher positive attitudes to robots tend to have a higher propensity to trust them. This result signifies the importance of gathering information on users' PTT and prior attitudes in robots when investigating human-robot trust, as this can help in understanding the processes underlying trust interactions, and hence, inform appropriate robot design to context [17].

In general as observable from Section V-E, subjects preferred the Extrovert Robot voice although the Introvert Robot speech was found to be clearer. Voice was also the main determinant for preference and trust. This is mainly due to the cheerful, friendly and appealing traits of the Extrovert Robot's voice and calm delivery and clarity of the Introvert Robot's speech. Both these qualities inspired trust, suggesting the importance of voice characteristics in moderating user trust. Additionally, the perceived competency, trustworthiness and affective influence of the robot are other factors affecting trust perception.

VIII. CONCLUSIONS AND FUTURE WORK

This paper investigates how users' individual differences affect trust in the Robo-Baristas. Although providing the first

step to investigate HRI for the service industry, it has some limitations that provide future research opportunities.

Our study revealed that subjects' predispositions (i.e. propensity to trust in robots and negative attitudes towards robots) predicted their levels of trust in the Robo-Baristas. Furthermore, to a certain extent subjects' Big 5 personality traits influenced their PTT and perceived trustworthiness of the Robo-Baristas. The main implication is that any human-robot trust study should consider collecting this information as it can inform context-appropriate robot design. Overall, subjects preferred and trusted an extrovert Robo-Barista more than an introvert Robo-Barista, irrespective of their own personality, suggesting the importance of context as well as prior stereotypes. Another implication of our findings is that voice features should be taken account when designing social robots and the resulting voice should sound trustworthy relative to the robot's role. Impression of competence and trustworthiness are other factors to consider to prevent misalignment of users' expectations and robot capabilities, which might lead to trust breakdown.

It would be interesting in future work to investigate not only subjects' personality but subjects' stereotypical thoughts on the task and the associated personality. Furthermore, instead of observing video vignettes of the interactions, in-person human-robot interactions where subjects order drinks from the Robo-Baristas is a direction worth exploring, as the relationship of personality with trust and performance in HRI may be moderated by the type of study [51]. Given that mental models of robots are context-sensitive, in future subjects should be given an option to flag out non-applicable items in the trust scale used [47] and these ratings should be considered when reporting and interpreting results. Cultural factors also play a role in how individuals trust robotic agents [17], providing another interesting research avenue. Finally, with trust being a dynamic process [9] and social robots being built to support long-term interactions, we are currently planning a long-term "in the wild" study to investigate trust formation and development over time in the café robot context.

REFERENCES

- [1] M. Y. Lim, J. David Aguas Lopes, D. A. Robb, B. W. Wilson, M. M., and H. Hastie, "Demonstration of a robo-barista for in the wild interactions," in *Proc. ACM/IEEE International Conference on Human-Robot Interaction (HRI'22)*, Sapporo, Hokaido, Japan, 2022.
- [2] O. L. Meriam Moujahid, Helen Hastie, "Multi-party interaction with a robot receptionist," in *Proc. of the ACM/IEEE International Conference on Human-Robot Interaction*, 2022.
- [3] S. Keizer, M. E. Foster, Z. Wang, and O. Lemon, "Machine learning for social multiparty human-robot interaction," *ACM Trans. on Interactive Intell. Syst.*, vol. 4, no. 14, pp. 1–32, 2014.
- [4] K. T. Wynne and J. B. Lyons, "An integrative model of autonomous agent teammate likeness," *Theor. Issues in Ergonom. Sci.*, vol. 19, pp. 353–374, 2018.
- [5] P. A. Hancock, T. T. Kessler, A. D. Kaplan, J. C. Brill, and J. L. Szalma, "Evolving trust in robots: Specification through sequential and comparative meta-analyses," *Human factors*, vol. 63, no. 7, pp. 1196–1229, 2021.
- [6] M. Jooisse, M. Lohse, J. G. Pérez, and V. Evers, "What you do is who you are: The role of task context in perceived social robot personality," in *Proc. of Robotics and automation 2013 IEEE international conference*. Karlsruhe, Germany: IEEE, May 2013, pp. 2134–2139.

- [7] E. Howarth, "Expectations concerning occupations in relation to extraversion-introversion," *Psychological Reports*, vol. 24, no. 2, pp. 415–418, Nov. 1999.
- [8] G. M. Sandstrom and E. W. Dunn, "Is efficiency overrated?: Minimal social interactions lead to belonging and positive affect," *Social Psychological and Personality Sci.*, vol. 5, no. 4, pp. 437–442, 2014.
- [9] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors: The J. of the Human Factors and Ergonom. Soc.*, vol. 46, pp. 50–80, 2004.
- [10] A. Freedy, E. DeVisser, G. Weltman, and N. Coeyman, "Measurement of trust in human–robot collaboration," in *Proc. of International symposium on collaborative technologies and systems*, 2007, pp. 106–114.
- [11] J. M. Kraus, "Psychological processes in the formation and calibration of trust in automation," Ph.D. dissertation, Ulm University, Germany, 2020.
- [12] T. Nomura, T. Suzuki, T. Kanda, and K. Kato, "Measurement of negative attitudes toward robots," *Interact. Studies: Social Behav. and Commun. in Biol. and Artif. Syst.*, vol. 7, no. 3, pp. 437–454, 2006.
- [13] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Acad. of Manage. Rev.*, vol. 20, pp. 709–734, 1995.
- [14] D. H. McKnight, L. L. Cummings, and N. L. Chervany, "Initial trust formation in new organizational relationships," *Acad. of Manage. Rev.*, vol. 23, no. 3, pp. 473–490, 1998.
- [15] J. B. Rotter, "A new scale for the measurement of interpersonal trust," *J. of personality*, vol. 35, no. 4, pp. 651–665, 1967.
- [16] S. A. Jessup, T. Schneider, G. Alarcon, T. Ryan, and A. Capiola, "The measurement of the propensity to trust automation," in *Virtual, Augmented and Mixed Reality. Applications and Case Studies. HCII 2019. Lecture Notes in Computer Science*, C. J. and F. G., Eds. Cham: Springer, 2019, vol. 11575.
- [17] M. Lewis, K. Sycara, and P. Walker, "The role of trust in human-robot interaction," *Found. of Trusted Autonomy*, 2018.
- [18] R. R. McCrae, "Human nature and culture: A trait perspective," *J. of Res. in Personality*, vol. 38, pp. 3–14, 2004.
- [19] S. O. Gaines, A. T. Panter, M. D. Lyde, W. N. Steers, C. E. Rusbult, C. L. Cox, and M. O. Wexler, "Evaluating the circumplexity of interpersonal traits and the manifestation of interpersonal traits in interpersonal trust," *J. of Personality and Social Psychol.*, vol. 73, no. 3, pp. 610–623, 1997.
- [20] S. M. Merritt and D. R. Ilgen, "Not all trust is created equal: Dispositional and history-based trust in human-automation interactions," *Human Factors*, vol. 50, no. 2, pp. 194–210, 2008.
- [21] A. M. Evans and W. Revelle, "Survey and behavioral measurements of interpersonal trust," *J. of Res. in Personality*, vol. 42, p. 1585–1593, 2008.
- [22] L. Robert, "Personality in the human robot interaction literature: A review and brief critique," in *Proc. of the 24th Americas conference on information systems*, L. P. Robert, Ed., New Orleans, LA, Aug. 2018, pp. 16–18.
- [23] M. Schmitz, A. Krüger, and S. Schmidt, "Modeling personality in voice of talking products through prosodic parameters," in *Proc. of IUI 2007– 12th International Conference on Intelligent User Interfaces*, Honolulu, Hawaii, USA, Jan 2007, pp. 313–316.
- [24] A. Tapus, C. Țăpuș, and M. Mataric, "User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy," *Intell. Service Robot.*, vol. 1, no. 2, pp. 169–183, 2008.
- [25] J. Pittam, *Voice in social interaction: An interdisciplinary approach*. Thousand Oaks, CA: Sage, 1994.
- [26] K. J. Tusing and J. P. Dillard, "The sounds of dominance: Vocal precursors of perceived dominance during interpersonal influence," *Human Commun. Res.*, vol. 26, no. 1, pp. 148–171, 2000.
- [27] F. Mairesse and M. Walker, "PERSONAGE: Personality generation for dialogue," in *Proc. of the 45th Annual Meeting of the Association of Computational Linguistics*. Prague, Czech Republic: Association for Computational Linguistics, Jun. 2007, pp. 496–503.
- [28] M. Neff, Y. Wang, R. Abbott, and M. Walker, "Evaluating the effect of gesture and language on personality perception in conversational agents," in *Proc. IVA 2010, Lecture Notes in Computer Science*, J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova, Eds. Berlin, Heidelberg: Springer, 2010.
- [29] A. Furnham, *Handbook of Language and Social Psychology*. Winley, 1990.
- [30] M. L. et al., "Evaluating extrovert and introvert behaviour of a domestic robot — a video study," in *Proc. of The 17th IEEE International Symposium on Robot and Human Interactive Communication*, Munich, 2008, pp. 488–493.
- [31] K. Leuwerink, "A robot with personality: Interacting with a group of humans," in *Proc. of the 16th Twente Student Conference on IT*, vol. 4, Enschede, The Netherlands, 2012.
- [32] G. Alarcon, J. Lyons, J. Christensen, M. Bowers, S. Klosterman, and A. Capiola, "The role of propensity to trust and the five factor model across the trust process," *J. Res. Pers.*, vol. 75, pp. 69–82, 2018.
- [33] S. Whittaker, Y. Rogers, E. Petrovskaya, and H. Zhuang, "Designing personas for expressive robots: Personality in the new breed of moving, speaking, and colorful social home robots," *ACM Trans. Human-Robot Interact.*, vol. 10, no. 8, p. 25 pages, February 2021.
- [34] K. M. Lee, W. Peng, S.-A. Jin, and C. Yan, "Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction," *J. of Commun.*, vol. 56, pp. 754–772, 2006.
- [35] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer, "Not so different after all: A cross-discipline view of trust," *Acad. of Manage. Rev.*, vol. 23, no. 3, pp. 393–404, 1998.
- [36] A. Mileounis, R. H. Cuijpers, and E. I. Barakova, "Creating robots with personality: The effect of personality on social intelligence," in *Artificial Computation in Biology and Medicine. IWINAC 2015. Lecture Notes in Computer Science*, F. V. J., Álvarez Sánchez J., de la Paz López F., T.-M. F., and A. H., Eds. Cham: Springer, 2015, vol. 9107.
- [37] J. Lopes and H. Hastie, "The language of persuasion, negotiation and trust," in *Proc. of Semdial*, 2021.
- [38] M. Rheu, J. Y. Shin, W. Peng, and J. Huh-Yoo, "Systematic review: Trust-building factors and implications for conversational agent design," *Int. J. of Human–Computer Interact.*, vol. 37, no. 1, pp. 81–96, 2021.
- [39] T. Bunk, D. Varshneya, V. Vlasov, and A. Nichol, "Diet: Lightweight language understanding for dialogue systems," *arXiv preprint arXiv:2004.09936*, 2020.
- [40] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. of NAACL*, Jun. 2019, pp. 4171–4186.
- [41] V. Vlasov, J. E. Mosig, and A. Nichol, "Dialogue transformers," *arXiv preprint arXiv:1910.00486*, 2019.
- [42] A. Mehrabian and J. A. Russell, *An approach to environmental psychology*. the MIT Press, 1974.
- [43] S. Srivastava. (2012) Norms for the big five inventory and other personality measures. [Online]. Available: <https://thehardestscience.com/2012/10/17/norms-for-the-big-five-inventory-and-other-personality-measures/>
- [44] J.-Y. Jian, A. Bisantz, and C. Drury, "Foundations for an empirically determined scale of trust in automated systems," *Int. J. of Cogn. Ergonom.*, vol. 4, pp. 53–71, 2000.
- [45] N. McDonald, S. Schoenebeck, and A. Forte., "Reliability and interrater reliability in qualitative research: Norms and guidelines for cscw and hci practice," in *Proc. ACM Human Computer Interaction*, no. 39, November 2019.
- [46] D. Johnson and K. Grayson, "Cognitive and affective trust in service relationships," *J. of Bus. Res.*, vol. 58, no. 4, pp. 500–507, 2005.
- [47] T. Law, N. Rabb, and M. Scheutz, "Can you trust your trust measure?" in *Proc. of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. New York, USA: ACM, March 2021, p. 9 pages.
- [48] A. Becker, T. Deckers, T. Dohmen, A. Falk, and F. Kosse, "The relationship between economic preferences and psychological personality measures," *Annual Review of Econ.*, vol. 4, pp. 453–478, 2012.
- [49] C. Calhoun, P. Bobko, M. Schuelke, S. Jessup, T. Ryan, C. Walter, L. Hirshfield, N. Bowling, C. Bragg, S. Khazon, A. Nelson, G. Alarcon, and C. Stokes, "Suspicion, trust, and automation," Tech. Rep. SRA Int. Inc. Publication No. AFRL-RH-WP-TR-2017-0002, 2017.
- [50] T. Dohmen, A. Falk, D. Huffman, and U. Sunde, "Representative trust and reciprocity: Prevalence and determinants," *Econ. Inquiry*, vol. 46, pp. 84–90, 2008.
- [51] G. M. Alarcon, A. Capiola, and M. D. Pfahler, "Chapter 7 - the role of human personality on trust in human-robot interaction," in *Trust in Human-Robot Interaction*, J. B. L. Chang S. Nam, Ed. Academic Press, 2021, pp. 159–178.