MUNI FSS

JOURNAL OF PSYCHOSOCIAL RESEARCH ON CYBERSPACE
CYBERPSYCHOLOGY

Tschopp, M., Gieselmann, M., & Sassenberg, K. (2023). Servant by default? How humans perceive their relationship with conversational AI. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 17*(3), Article 9. https://doi.org/10.5817/CP2023-3-9

Servant by Default? How Humans Perceive Their Relationship With Conversational AI

Marisa Tschopp[1], Miriam Gieselmann[2], & Kai Sassenberg[2,3]

[1] Scip AG, Zürich, Switzerland
[2] Leibniz Institut für Wissensmedien (IWM), Tübingen, Germany
[3] University of Tübingen, Tübingen, Germany

Abstract

Conversational AI, like Amazon's Alexa, are often marketed as tools assisting owners, but humans anthropomorphize computers, suggesting that they bond with their devices beyond an owner-tool relationship. Little empirical research has studied human-AI relationships besides relational proxies such as trust. We explored the relationships people form with conversational AI based on the Relational Models Theory (RMT, Fiske, 1992). Results of the factor analyses among frequent users (N_{total} = 729) suggest that they perceive the relationship more as a master-assistant relationship (i.e., authority ranking) and an exchange relationship (i.e., market pricing) than as a companion-like relationship (i.e., peer bonding). The correlational analysis showed that authority ranking barely correlates with system perception or user characteristics, whereas market pricing and peer bonding do. The relationship perception proved to be independent of demographic factors and label of the digital device. Our research enriches the traditional dichotomous approach. The extent to which users see their conversational AI as exchange partners or peer-like has a stronger predictive value regarding human-like system perception of conversational AI than the perception of it as servants.

Keywords: conversational AI; voice assistant; human-AI relationship; system perception; user characteristics

Editorial Record

First submission received:
May 3, 2022

Revisions received:
November 22, 2022
March 14, 2023
May 23, 2023

Accepted for publication:
May 26, 2023

Editor in charge:
David Smahel

Introduction

"Alexa, will you marry me?" 6,000 times per day users propose to Alexa and 19,000 times per day, users in India say "I love you" to Alexa, Amazon's conversational AI (Amazon, 2021)[1]. Furthermore, 14% of male Alexa users in the UK desire a sexual relationship with "her" (The Guardian, 2020), and 3.5 million times Germans said "I love you" in the first half of 2021 (Buschke, 2021). But just because users say they "love" Alexa does not necessarily mean they are in love with "her" as humans are in love with other humans. So, if people do not really love Alexa but also do not see Alexa as "just" a tool, how do they actually perceive Alexa in relation to themselves?

The aforementioned numbers are congruent with the ample corpus of research informed by the computers as social actors (CASA; Nass & Moon, 2000; Nass et al., 1994) paradigm, suggesting that humans apply social rules from human interaction when they communicate with machines (Gambino et al., 2020). There is still a tension between the fact that "Alexa is just a tool" and the anecdotal evidence above suggesting that humans might as well build more equal or even peer-like relationships with their conversational AI, which is why relationship research around conversational AI has been repeatedly called for (Guzman & Lewis, 2020; Hepp, 2020; Kim et al., 2019; Seymore & van Kleek, 2021).

Whereas some AI systems are intentionally designed to fulfill a single role, e.g., a friend or lover, see for instance, the chatbot Replika (Skjuve et al., 2021), it is not entirely clear what off-the-shelf conversational AI, such as Alexa, are designed for (Purington et al., 2017). A friend? An assistant? Or both?

Thus, the present study aims to address this gap by investigating how humans perceive their relationship to conversational AI. To this end, we apply the Relational Models Theory (RMT; Fiske, 1992) to human-AI relationships. Fiske suggested a comprehensive framework of relationships between humans consisting of four different modes of relationships that could also inform our understanding of how people perceive their relationship with conversational AI.

By using Fiske's framework, we aim to take a more differentiated look at the perceived relationships going beyond simple dichotomous or two-dimensional approaches commonly used in earlier research, for instance, comparing servant (transactional) and friend (relational) roles (Kim et al., 2019). A better understanding of the relationship humans perceive with their conversational AI will inform not only the interface design of these devices but also open avenues for new research questions: What relational elements are more relevant, for instance, when turning on the lights versus for voice shopping or when sharing personal information with one's device? Being the first to apply RMT to human-AI interaction, we aim to contribute to research and practice by using a well-established, quantitative framework that mirrors increasing reciprocal dynamics rather than stagnant role ascriptions. Overall, we aim to provide further empirical evidence for the overarching question: In what respects are digital assistants more than just tools?

State of the Art of Human-AI Relationship Research

Unlike human relationships, the relation of a human to a machine is unidirectional because a device is not a living being. However, the growing capabilities in natural language understanding and processing (Guzman, 2019; Panetta, 2020; van Berkel et al., 2021) could obscure unidirectionality and affect how people perceive themselves in relation to the machine. What is currently missing is a theoretical and quantitative measurement approach that mirrors the increasing reciprocal character and offers a variety of roles users ascribe to their conversational AI.

The first approaches to studying the social dynamics in human-AI interaction drew on findings from neighboring contexts, such as social robotics (Sundar, 2020; van Berkel et al., 2021). Some studies relied on the theories of role categorizations (e.g., robots as companions vs. assistants as in Sundar et al., 2017; based on Dautenhahn, 2007; or the social friend role as in Rhee & Choi, 2020) to investigate role ascriptions to conversational AI. In one study, Kim et al. (2019) found that the relationship type "friend" led to higher perceptions of warmth (fully mediated by anthropomorphism) compared to a "servant" relationship. Other researchers grounded their research in the social presence theory (i.e., the idea of feeling a deeper connection to a computer system, e.g., McLean & Osei-Frimpong, 2019): For example, Ki et al. (2020) focused on the role of para-friendships (i.e., humans perceiving emotional friend-like relationships, for instance, with fictional characters, or online influencers). They found that the perception of intimacy, understanding, enjoyability, and involvement affected para-friendships and fostered users' intention to use a digital assistant.

Building upon the CASA paradigm and humans' tendency to anthropomorphize (Epley et al, 2007; Waytz et al., 2010), a lot of studies in the broader field of human-AI interaction focused on the human-like perception of AI systems (assessing constructs such as perceived warmth, competence, or anthropomorphism, e.g., Bergmann et al., 2012; Gilad et al., 2021; Gong, 2008; Purington et al., 2017). Therefore, it is plausible to refer to social relations between humans to understand human-machine relations. Nonetheless, this approach has barely been taken.

For instance, in a survey study with Alexa users, Seymore and van Kleek (2021) used Knapp's (1978) staircase model. They found that the human-AI relationship development follows similar patterns as the relationship development among humans and that more developed relationships co-occurred with more anthropomorphizing of and more trust in the conversational AI. In addition, two qualitative studies on the companion chatbot Replika relying on different attachment theories (e.g., social penetration theory by Taylor & Altman, 1987, was used by Skjuve et al., 2021) conclude that theories about interpersonal relationships can be used to gain further insights into the development of human-AI relationships (see also Xie & Pentina, 2022).

The studies mentioned above indicate that users are developing relationships with AI systems similar to those with humans. They offer insights informative for specific goals; for instance, design recommendations to increase adoption. Furthermore, these qualitative studies offer rich insights into the drivers of human-computer relationships (Skjuve et al., 2021; Xie & Pentina, 2022). On the other hand, recent studies by Lopatovska and Williams (2018) and Croes and Anthenuis (2021) found no clear evidence for humans building emotional relationships with conversational AI. For an overview of recent studies, see Table 1.

Table 1. *Overview of Related Research on Human-AI Relationships.*

Theoretical Basis	Author(s), Year	Methods	Key Findings
Role categorizations (Dautenhahn, 2007)	1) Sundar et al., 2017 2) Kim et al., 2019 3) Hu et al., 2022 4) Rhee & Choi, 2020	Quantitative	1) Congruity of role and demeanor matters 2) Friend role positively related to warmth perception 3) Power over AI (servant role of conversational AI) reduces risk perception in voice shopping 4) Friend-role of conversational AI influences positive attitude toward a product
Social presence theory (Short et al., 1976), para-social relationships (Horton & Wohl, 1956)	Ki et al., 2020; McLean and Osei-Frimpong, 2019	Quantitative	Para-friendships / social presence foster usage intention
Personification, anthropomorphism, CASA paradigm, relationship indicators (Epley et al., 2007; Nass & Moon, 2000)	1) Lopatovska and Williams, 2018 2) Schweitzer et al., 2019 3) Purington et al., 2017 4) Gong, 2008 5) Gilad et al., 2021 6) Bergmann et al., 2012	1), 2) Qualitative 3), 4), 5), 6) Quantitative	1) Relationship indicators are rather mindless 2) Friend, servant, or master relationship types may develop 3) Personification linked to more social interaction 4) More anthropomorphic agent received more social responses 5) Users prefer high-warmth systems 6) Warmth/competence perception depend on time, agent appearance/behavior
Knapp's staircase model (Knapp, 1987)	Seymore and van Kleek, 2021	Quantitative	Users show attachment to their conversational AI
Social penetration theory (Taylor & Altman, 1987)	1) Croes and Anthenuis., 2021 2) Xu and Li, 2022 3) Skjuve et al., 2021	1), 2) Quantitative 3) Qualitative	1) No indicators of a progressing human-AI relationship similar to humans 2) Functional and relational use influence each other 3) Indicators of a progressing human-AI relationship similar to humans
Bowlby's attachment theory (Bowlby, 1979)	1) Xie and Pentina, 2022 2) Pentina et al., 2023	1) Qualitative 2) Mixed Methods	1) Users developed emotional bonds with their companion 2) Identification of various antecedents plus moderators that influence attachment to conversational AI

We argue that one problem is that many studies are, apart from the qualitative studies, arguably insufficient to comprehensively examine human-AI relationships, especially for off-the-shelf conversational AI, which serve multiple purposes. Ki et al. (2020) considered exclusively (para-)friendship, Seymore and van Kleek (2021) focused on relationship development, and Kim et al. (2019) focused on a dichotomous distinction of the friend or servant as relationship types. Exchange-oriented relationships have, for instance, not been considered so far. Along these lines, qualitative evidence from a study on conversational AI users indicated that participants build a diversity of relationships with their devices, suggesting that a more systematic, quantitative study of human-AI relationship perception is necessary (Schweitzer et al., 2019).

Therefore, we aim to build upon these findings by applying a comprehensive relationship model from human-to-human relations to the human-AI relationship that goes beyond one or only two relational types studied in earlier work by considering multiple independent relational dimensions. To the best of our knowledge, no studies take such a multidimensional approach to the perceived relationship between humans and commercially available conversational AI. Building upon the strongly established RMT (Haslam & Fiske, 1999) seems a promising direction to understand user perceptions of human-AI relationships as a whole, also compared to other approaches capturing single relational proxies, like trust or perceived warmth.

Relationships Between Humans According to Fiske (1992)

Fiske (1992) proposed and received mighty empirical support (Haslam & Fiske, 1999) for the RMT proposing four modes or dimensions along which humans perceive their relationships to other humans: authority ranking, market pricing, communal sharing, and equality matching.

Authority ranking denotes an asymmetric relationship where people are not equivalent but ordered along some hierarchical dimension, where the highest rank is entitled to command over and protect the lower ranks (e.g., commander and soldier, parent and child). *Market pricing* is all about rational cost-benefit analysis (i.e., a relationship mode focusing on exchange), where humans seek something in return for their investment in the relationship, for instance, money (e.g., people in working groups). *Communal sharing* is high in tribes with kinship-like relations, where people are equivalent and commonalities are emphasized (e.g., often in families or sports teams). *Equality matching* describes a peer-like tit-for-tat relationship, focusing on an egalitarian balance within the relationship (e.g., roommates sharing an apartment). According to RMT, the four modes are not distinct types but can instead operate at the same time—they are dimensions along which one relationship with a specific person can be described. A brief overview of the human relationship modes is depicted in Table 2.

Table 2. *Brief Description of the Four Modes of Relationships According to Fiske (1992).*

Authority Ranking	Market Pricing	Communal Sharing	Equality Matching
What is the order between us?	What are the ratios?	What do we have in common?	What is the balance?
Hierarchy	Cost-benefit	Solidarity	Turn-taking
Command	Input-Output	In-group	Democratic voting
Dominance	Reciprocity	Communality	Reciprocity

Applying Relational Models Theory (RMT) to Human-AI Relationships

How could the human relationship modes apply to human-AI relationships?

The presentation of conversational AI as digital assistants is in line with a strong emphasis on the *authority ranking* mode. This would be a rational perception of the human-AI relationship: the user dominates the conversational AI and gives commands. Therefore, authority ranking might be the dominant relational mode—in other words, the one with the highest mean values.

The actual and perceived growth of machine agency (i.e., machines and algorithms are becoming more agent-like, having more control, and exerting greater influence on people's perceptions and behavior; Sundar, 2020), provides the basis for perceived reciprocity. Accordingly, the relationship to conversational AI should, in part, be seen as characterized by market pricing. However, nowadays, conversational AI does not come close enough to human characteristics that users will likely experience solidarity and community—attributes that allow for communal sharing or actual perceived equality—the precondition for equality matching. Therefore, we predict:

H1: Authority ranking and market pricing have significantly higher mean values than equality matching and communal sharing.[2]

Various streams of human-machine interaction studies follow the dichotomous distinction between rational and emotional dimensions of how humans perceive and interact with machines (for a detailed discussion, see Glikson & Woolley, 2020). We follow this sound distinction when deriving predictions for concepts correlating with the dimensions of the human-AI relationships derived from RMT.

Rational Modes of Perceived Human-AI Relationships

The extent to which people perceive their relationship with a conversational AI to be characterized by authority ranking and market pricing, should be related to rational concepts of system perception, like the perceived competence of a system. Perceived competence is a prerequisite for the assistant role as well as for perceiving the conversational AI as an equal counterpart as implied by marketing pricing. Conversely, both relationship dimensions should be negatively related to competence concerns. For authority ranking this has been shown by Hu et al. (2022). They found that more perceived power over Alexa (Amazon's conversational AI) led to reduced risk perception—with regard to shopping at Amazon (provided that the people had a desire for power). Together, this led to the following prediction:

H2: Authority ranking and market pricing are related to rational dimensions of system perception—positively to perceived competence and negatively to perceived competence concerns.

Emotional Modes of Perceived Human-AI Relationships

As the examples in the introduction indicate, people might form emotional relations to conversational AI of varying depth. In human relationships, trust is one of the key ingredients in communal sharing. To a lesser extent, also in equality matching (Fiske, 1992), and indeed users who categorize their AI systems as friends reported a higher level of trust, perceived warmth, and lower social distance (Bergmann et al., 2012; Gilad et al., 2021; Pitardi & Marriot, 2021). They would likely perceive their relationship as more emotional; hence, the communal relationship mode would likely occur in human-AI interaction.

People who experience an emotional relationship with a machine assume that it is characterized by similar feelings and behaviors to human friends (e.g., Han & Yang, 2018), and attribute constructs such as low psychological distance (Pitardi & Marriot, 2021) to their device. Seymor and van Kleek (2021) found a positive correlation between trust, anthropomorphism, and the closeness of the relationship (see also Schweitzer et al., 2019). Hence, communal sharing and equality matching are likely related to more anthropomorphizing of the conversational AI, a higher perceived warmth, lower perceived social distance, and higher trust. These considerations led us to the following prediction:

H3: Equality matching and communal sharing are positively related to emotional dimensions of system perception: perceived warmth, inclusion of AI in self, psychological distance, anthropomorphism, and trust.

The Current Research

To enhance the existing literature, which largely relies on one or two-dimensional frameworks to study perceptions of relationships, we repurposed the empirically established multidimensional relational models framework (RMT) by Fiske (1992) to conversational AI in two studies (N_1 = 367 and N_2 = 362). We aim to test the three hypotheses derived above in Study 1 by collecting variables of system perception: system perception comprised variables of a rather rational nature, namely, perceived competence, competence concerns, and privacy concerns. Furthermore, variables of a rather emotional nature, namely, perceived warmth, inclusion of self in AI, psychological distance, and anthropomorphism, were collected. We measured trust under system perception, with the peculiar feature that human trust in AI shares both emotional as well as rational properties (see Glikson & Woolley, 2020; Jian et al., 2000). For additional analyses, we collected user characteristics variables, namely, affinity to technology, frequency of use, years of experience with the conversational AI, different purposes of usage, as well as the number of purposes, which we derived from the sum of provided answers of participants for purposes of use (e.g., navigation, shopping, or controlling the smart home).

In Study 2, we aimed to replicate the dimensional structure and the prevalence of the relationship modes (H1). In addition, we sought to test the stability of the relationship modes and their prevalence across a number of user characteristics. To this end, we assessed a variety of demographics (as suggested by McLean & Osei-Frimpong, 2019 and Purington et al., 2017; i.e., household specifications, educational level, employment, age, gender of the user, and technological knowledge) and system variables (i.e., device specifications, set gender of the technology). In addition, we tested the preregistered hypothesis (https://aspredicted.org/qa5ty.pdf) that naming the device "voice assistant" (compared to "conversational AI") will lead to higher perceived authority ranking and lower perceived peer-bonding as well as market pricing (H4). This is because using the word assistant in the label will stress the hierarchical relationship captured by authority ranking[3]. Not finding such an effect and no relevant correlations with the demographics and system characteristics will be a sign that the relationship perceptions are stable and largely independent of personal and situational factors.

Study 1

Methods

Participants

We conducted a correlational online questionnaire study via Prolific in June 2021. In a screening study (N_0 = 1,050), we surveyed participants for regular use of conversational AI, such as Alexa, Siri, or Google Assistant. Our target sample size was 400. To ensure stable estimates of correlations, Schönbrodt and Perugini (2013) recommend a sample size of around 250. To account for potential exclusions, we added 150 observations to our sample. Of those who reported using a conversational AI regularly, we invited N_1 = 406 participants to participate in study 1 in exchange for £1.25. We excluded 39 participants from the analyses reported below who failed at least one of two attention checks. Participants whose duration of taking the survey was too short for reading the items properly (< 150 seconds) or excessively long (> 80,000 seconds) were also excluded. The three hundred sixty-seven (N = 367) remaining respondents (67% female, 33% male; age range: 18–60 years; 68% between 25 and 49 years) were mostly from the UK (89%) and filled out the questionnaire on a mobile device (44%), desktop computer (52%), or tablet (4%). The majority possessed multiple conversational AI: Alexa (43%), Siri (25%), or Google Assistant (24 %), and less than 2% used Cortana, Bixby, or other conversational AI (multiple responses were possible).

Procedure

We invited participants to participate in a study on users' perceptions of technology. After providing consent, participants first answered questions about their usage of conversational AI. Then the perception of their human-AI relationship (adapted from MORQ, Haslam & Fiske, 1999). The instructions for the original measure require people to focus on a counterpart to describe their relationship, referring *only* to that counterpart. In line with this approach, we asked our respondents to focus on one specific conversational AI while answering the questions regarding the human-AI relationship. Most of the respondents chose Alexa (63.5%), the Google Assistant (23.2%), and Siri (12%, 1.3% other) Afterward, we collected the other scales to measure system perception in a fixed order (psychological distance, social perception, inclusion of others, trust, and anthropomorphism), items of each scale were presented in randomized order. Then we collected data on user characteristics, including affinity to technology, frequency, experience, and purposes of use. Finally, we collected barriers to usage, particularly privacy and competence concerns. We placed demographics at the end, followed by a final opportunity to withdraw the data before submission.

To check for common method bias, we looked at Harman`s Single factor test to check the extent of common method bias, which indicated a shared variance of 11.9% for all measures reported in Table 3; similarly, the latent common method factor approach indicated a shared variance of 15.5% for all measures including more than two items, indicating an acceptable level of shared variance across measures.

Measures

Human-AI Relationship. We assessed the human-AI relationship with a version of the MORQ (Modes of Relations Questionnaire, Haslam & Fiske, 1999) adapted to the context of conversational AI. As this questionnaire was repurposed for the first time, we first carefully selected items we deemed appropriate for the context of human-AI interaction. We aimed to stick as closely as possible to the original questionnaire but had to omit items that we considered not applicable to conversational AI, for instance, because they refer to physical objects such as *You typically divide things up into shares that are the same size*. We pretested the remaining items with fifteen people with expertise in IT, resulting in adaptions in the wording to render it suitable for human-AI relationships. The final questionnaire consisted of 17 items (see Table 2) using a 7-point Likert scale (1 = *not at all true for this relationship*, 7 = *very true for this relationship*): Five items for communal sharing, four items for equality matching, four items for authority ranking, and four items for market pricing. Given that this instrument was used for the first time in such a context, we report detailed analyses of the dimensional structure and internal consistencies in the Results section.

Trust. Trust was measured through the Trust in Automation Scale (Jian et al., 2000). We slightly adapted the wording (conversational AI instead of "the automation") of the 13 items (Cronbach's Alpha = .78), where respondents had to indicate how much they agree to a statement on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). For example, *The cAI is dependable* (in the beginning, we explained that conversational AI is abbreviated cAI). Here and for all other scales, an index was formed by averaging the responses after recoding reversed items.

Social Perception. Social perception (perceived warmth and competence) was measured with six items on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*) slightly adapted from Pitardi and Marriot (2021), where three items measured the perceived warmth dimension (e.g., *I think the cAI has good intentions*) and the other three items measured the competence dimension (e.g., *I think the cAI is effective*). The factor analysis across these six items revealed a two-factor structure, with two items assigned to the perceived warmth dimension and four items assigned to the competence dimension. The final scales presented decent reliabilities with Cronbach's Alpha = .81 for the competence dimension and r = .48 (N = 367, p < .001) for the perceived warmth dimension.

Inclusion of AI in the Self. Inclusion of AI in the self was measured with two items on a 7-point Likert scale (r = .61, N = 367, p < .001). The items are two pictorial measures of categorization which were adapted from Schubert and Otten (2002; Aron & Smollan, 1992). The first item consisted of seven pictures of two equivalent circles on a straight line. As in the original, the circles moved closer together from the top to the bottom, overlapping completely in the last bottom picture. The instructions indicated that one circle represented the "self" or humans in general (differing between the two items) and the other the "conversational AI". The higher the score, the higher the overlap, and the closer the respondent felt to the conversational AI.

Psychological Distance. Psychological distance was measured with two items (r = .77, N = 367, p < .001) on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*). We adapted the items slightly, for instance, *the cAI is psychologically close to me*, and dropped the item measuring familiarity from the original (Li & Sung, 2021).

Anthropomorphism. Anthropomorphism was measured through the scale by Waytz et al. (2010). Participants had to indicate their agreement with seven items (e.g.*, To what extent does the cAI have thoughts of its own?,* Cronbach's Alpha = .90) on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*).

Affinity to Technology. Affinity to technology was measured with nine items (e.g., *I like to occupy myself in greater detail with technical systems*, Cronbach's Alpha = .88) on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*) developed by Franke et al. (2019).

Frequency of Use. Frequency of use (How often do you use conversational AI?) was measured on a single-item 5-point scale ranging from 1 = *less than once a month* to 5 = *several times a day* (adapted from Funk et al., 2021).

Experience of Use. Experience of use (Since when do you use conversational AI?), was measured on a single-item 5-point scale ranging from 1 = *less than 12 months* to 5 = 5 *years or more*. This item was adapted from Funk et al. (2021).

Purpose of Use. The purpose of use was assessed to investigate for what purposes they use a conversational AI. Six answer options were given (multiple responses possible), e.g., option one was *To retrieve information, e.g., How is the weather tomorrow?* An open text option was given to collect purposes not captured within the preselected options. We calculated the *number of purposes* from the sum of retrieved answers from the participants (inspired by Funk et al., 2021).

Finally, we collected barriers of usage to address concerns users may have, where participants had to rate their agreement on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). We conducted a factor analysis across these seven items, which revealed a two-factor structure. We distinguished privacy concerns (e.g., *I am concerned about privacy*, r = .90, N = 367, p < .001, 2 items) and competence concerns (e.g., *The conversational AI cannot do what I expected it to do*, Cronbach's Alpha = .75, 5 items). Items were adapted from Funk et al. (2021).

Analyses have been conducted using SPSS 25.0 unless reported otherwise. Higher values indicate a stronger manifestation of the respective construct for all scales. Code, data, codebook, and calculations in excel for both studies are available here: https://researchbox.org/636

Results

Factor Structure of Human-AI Relationship Questionnaire

A principal component analysis (PCA) with orthogonal rotation (varimax) was conducted on the 17 items of the human-AI relationship measure. The Kaiser criterion as well as the visual inspection of the scree plot indicated that a three-component solution is adequate and explained 53.8% of the variance. Table 3 shows the factor loadings after rotation. The first component combines the two original modes—communal sharing and equality matching. It, thus represents an emotional, peer-like character of human-AI relationships, which we will name *peer bonding* hereafter. Analogous to the original questionnaire, component 2 represents authority ranking and component 3 represents market pricing. We omitted Item 7 from the analyses because it had the highest loading on a different factor than in the original scale. No other items were omitted, even though some had meaningful secondary loadings, to keep the scales as close as possible to MORQ.

Table 3. *Results From a Factor Analysis of the Human-AI Relationship Questionnaire (N_1 = 376).*

	Item	Factor Loading		
		1	2	3
Communal sharing				
1	There is a moral obligation to act kindly to each other	**.474**		.370
2	Decisions are made together	**.815**		
3	You tend to develop similar attitudes and behaviors	**.788**		
4	It seems you have something unique in common	**.845**		
5	The two of you belong together	**.714**		
Equality matching				
6	Some requests are granted in anticipation of something in return	**.557**		
7	"One-Person, one vote"' is the principle for making decisions	.387	**.505**	
8	You take turns doing what the other wants.	**.761**		
9	You are like peers or fellow co-partners	**.738**		
Authority ranking				
10	One of us is entitled to more than the other		**.722**	
11	One directs the work, the other pretty much follows		**.554**	.459
12	You are like leader and follower		**.698**	
13	One is above the other in a kind of hierarchy		**.751**	
Market pricing				
14	What you get is directly proportional to how much you give	.311		**.616**
15	You have a right to a fair rate of return for what you put into this interaction			**.741**
16	You expect the same return on your investment other people get			**.741**
17	Your interaction is a strictly rational cost-benefit analysis		.404	**.420**

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 5 iterations. Highest factor loadings are in bold, factor loadings below .30 are not displayed. Adapted from the Modes of Relations Questionnaire by Haslam & Fiske, 1999.

Given the secondary loadings and the imperfect internal consistency for authority ranking, we conducted a confirmatory factor analysis using the R package "lavaan" (R version 4.1.2.; Rosseel et al., 2021). Model fit statistics indicated a satisfactory fit: $\chi^2(87) = 191$, $p < .001$; CFI = .94, RMSEA = .06, SRMR = .07.

The final scales presented decent reliabilities: Cronbach's Alpha = .86 for peer bonding (communal sharing and equality matching merged), Cronbach's Alpha = .70 for authority ranking, and Cronbach's Alpha = .68 for market pricing. The scales could not be substantially improved by removing items with the lowest item-total correlation.

To establish convergent validity, we calculated the average variance extracted (AVE) and composite reliability scores for each factor. The average variance extracted was .48 (factor 1), .46 (factor 2), and .61 (factor 3). Values for composite reliability ranged from .73 to .89, which exceeds .70 indicating that all the items consistently measured their corresponding constructs. To assess discriminant validity, we furthermore tested the Fornell-Larcker Criterion, where all conditions were satisfied, suggesting acceptable construct validity.

Market pricing was positively correlated with peer bonding ($r = .39$, $N = 367$, $p < .001$) and authority ranking ($r = .38$, $N = 367$, $p < .001$). No significant correlation was found between authority ranking and peer bonding ($r = -.03$, $N = 367$, $p = .512$).

To test Hypothesis 1, we conducted an ANOVA with repeated measures and post-hoc comparison using Bonferroni correction. As the sphericity assumption was violated, we report results with Huyn-Feldt correction. In line with the hypothesis, participants saw their relationship with the conversational AI as more strongly characterized by authority ranking ($M = 4.74$, $SD = 1.49$, $N = 367$) than by market pricing ($M = 4.29$, $SD = 1.33$, $N = 367$) and peer bonding ($M = 2.46$, $SD = 1.25$, $N = 367$), all ps < .001, $F(1.73, 632.30) = 378.63$, $p < .001$, $\eta_p^2 = .508$.

Relation Between Human-AI Relationships and Variables of System Perception and User Characteristics

In what follows, we report bivariate correlations and partial correlations of each relationship mode, controlling for the respective two other modes. We will focus more on the latter to test for unique relations between the respective relationship mode and the criterion variable. In doing so, we aim to address the hypotheses that the rather rational variables of system perception are related to authority ranking and market pricing (H2), whereas the emotional variables of system perception are related to equality matching and communal sharing (peer bonding, respectively; H3). In addition, we explored the relationship between user characteristics and relationship perception. Partial correlations, as well as bivariate correlations in parentheses, are reported in Table 4.

For authority ranking, we found only zero-order correlations with trust, perceived competence, affinity to technology, and the sum of purposes. After controlling for the other two relationship modes, these relations turned insignificant, except for the number of purposes. The more users perceived their relationship to a conversational AI to be characterized by authority ranking, the more widely they used it. Overall, authority ranking is the dominant mode regarding the perception of conversational AI, but the extent of authority ranking does not uniquely relate to the system perception. To our surprise, even the bivariate correlations with trust and perceived competence were small in size—though significant. This is somewhat peculiar given the role perceived competence and trust seem to play in human-AI interaction, or the way conversational AI are marketed as digital assistants.

Table 4. *Partial (Bivariate) Correlations Between Relational Modes (Controlled for the Respective Other Two in Case of the Partial Correlations) and System Perception as Well as User Characteristics (N = 376).*

Variable	Authority Ranking	Market Pricing	Peer Bonding
System perception			
Perceived competence	.05 (.13*)	.19** (.30**)	.15** (.25**)
Competence concerns	.04 (−.02)	−.14** (−.17**)	−.05 (−.12*)
Privacy concerns	−.03 (−.04)	−.01 (−.02)	.02 (.02)
Perceived warmth	.07 (.09)	.09 (.28**)	.37** (.43**)
Inclusion of AI in self	.03 (.05)	.12* (.33**)	.48** (.55**)
Psychological distance	−.01 (−.02)	.04 (.27**)	.59** (.64**)
Anthropomorphism	.01 (−.03)	−.05 (.16**)	.50** (.53**)
Trust	.04 (.12*)	.19** (.26**)	.05 (.13*)
User characteristics			
Affinity to technology	.07 (.12*)	.10* (.20**)	.11* (.16**)
Frequency of use	−.09 (−.07)	.05 (−.01)	−.08 (−.06)
Experience of use	−.08 (−.08)	.02 (−.05)	−.09 (−.09)
Number of purposes	.13* (.14*)	.01 (.12*)	.15** (.16**)

Note. Values in parentheses represent bivariate correlations. *$p < .05$. **$p < .01$.

Market pricing presented positive zero-order correlations with trust, warmth, inclusion of self in AI, psychological distance, anthropomorphism, perceived competence, affinity to technology, competence concerns (negative), and the number of purposes. However, market pricing had no unique relation (partial correlation) with warmth, psychological distance, anthropomorphism, and the number of purposes, whereas it had a unique relation with trust, inclusion of AI in the self, perceived competence, affinity to technology, and competence concerns (negative). Taken together, a higher market pricing mode was uniquely related to the rational dimensions (perceived competence and competence concerns) as well as to inclusion of AI in the self (emotional system perception) and trust, sharing features of both rational and emotional dimensions. Effects were small (to medium).

Peer bonding showed strong, positive bivariate and partial correlations with perceived warmth, inclusion of self in AI, psychological distance, and anthropomorphism. The same was true for perceived competence, affinity to technology, and the number of purposes. In addition, we found small, bivariate correlations with trust (positive) and competence concerns (negative), but the effects disappeared in the partial correlation analysis. In line with the idea that users build social bonds with computers, but contradicting the rational approach, peer bonding correlates strongest with system perceptions of conversational AI and also partly with user characteristics.

In a nutshell, we found support for H3, observing a stronger relationship between peer bonding and the emotional variables of system perception (e.g., psychological distance) rather than with the rational variables (e.g., perceived competence). With regards to H2, the results differ between market pricing and authority ranking. Market pricing showed positive correlations with rational and emotional variables of system perception, whereas authority ranking showed no partial correlations, which is not in line with our theorizing. Furthermore, no meaningful correlations were found between the user characteristics and the relational modes.

Discussion

The results of study 1 showed that participants view their relationship along three dimensions, instead of the suggested four dimensions. Equality matching and communal sharing merged into one dimension we called *peer bonding*. In line with our expectations (H1), we found that most users characterize their relationship as hierarchical (i.e., authority ranking) as well as the non-hierarchical exchange-based relationship (i.e., market pricing). Only a few saw their relationship characterized as a companion-like relationship (i.e., peer bonding). Against our expectations (H2), authority ranking did not show any meaningful correlations, rendering this traditional dimension not informative. Taking equality matching and communal sharing together, rather in line with our expectations (H3), peer bonding was to a greater extent related to the emotional variables of system perception, such as anthropomorphism or perceived warmth, than to the rational variables of system perception. Interestingly, trust was not related to peer bonding but market pricing, suggesting that users take a more rational approach that includes elements of exchange, which is in line with current critique on emotional accounts of trust (Ryan, 2020). In a nutshell, the relationship seems rather rational, but still relational. Finally, we observed no meaningful correlations with user characteristics, apart from the sum of purposes. Using conversational AI for multiple purposes (for instance, navigation, smart home, *and* voice shopping) may be a better indicator of "power users" than the simple frequency of use.

Given that this approach was used for the first time, we conducted a second study to replicate the factor structure and to test the stability of the hierarchy of the three dimensions considering demographic variables and a manipulation of the device's naming.

Study 2

Methods

Design, Participants, and Procedure

Study 2 was an experiment with two conditions varying the naming of the technology (voice assistant vs. conversational AI). Otherwise, Study 2 closely followed the procedure of study 2 but included a different set of measures apart from the relationship modes. We sampled 400 frequent users of conversational AI via Prolific in November 2022 (based on a prescreening). Similar to study 1, our rationale for determining the sample sized is based on the arguments by Schönbrodt and Perugini (2013). Excluding participants based on our preregistered exclusion criteria (https://aspredicted.org/qa5ty.pdf) would have led to an exclusion rate above 25%. Therefore, we applied a more lenient criterion than preregistered regarding the minimum time taken to complete the study (120 rather than 105 seconds). Notably, the results were not contingent on the exclusion criterion. The remaining

N = 362 respondents from study 2 (44% male, 56% female, age range 19–81 years, M_{age} = 39, SD_{age} = 13) were mostly from the UK (90%). 20% used their conversational AI on their smartphone, 61% on their smart speaker without a screen, and 15% on a smart speaker with a screen. This sample size was in line with the desired sample size determined by an a priori power analysis (2 (naming, between subjects factor) x 3 (relationship type, within subjects factor) mixed ANOVA, interaction, effect size f = .10, α = .05, 1–β = .95, N = 362). We aim for 95% power to be able to interpret non-significant results.

Measures

Human-AI Relationship. The perceived human-AI relationship was assessed using the scale from study 1 with minor adaptions in the wording. Detailed analyses of the dimensional structure and internal consistencies are reported in the Results section.

We collected the following variables to test their correlation to the relationship modes: household specifications single (0 = *no*,1 = *yes*), *kids* (0 = *no*,1 = *yes*), educational level (ranging from 1: *less than high school* to 7: *doctorate*), and employment (employed versus unemployed: 0 = *not employed*,1 = *employed*). In addition, we surveyed their technological knowledge on a 5-point Likert scale (1 = *not knowledgeable at all* to 5 = *extremely knowledgeable*), device details (without versus with screen: 0 = *without*, 1 = *with*), and the customized gender of the conversational AI (*female* = 1; *other* = 0).

Results

Factor Structure of Human-AI Relationship Questionnaire

The PCA with orthogonal rotation (varimax) on the 17 items of the human-AI relationship measure yielded similar results as in study 1. The Kaiser criterion as well as the visual inspection of the scree plot indicated that a three-component solution is adequate and explained 59.8% of the variance. We have omitted item 6 due to high secondary loading. Notably, including item 6 did not change results substantially.

The final scales presented decent reliabilities: Cronbach's Alpha = .90 for peer bonding (communal sharing and equality matching merged, 8 items), Cronbach's Alpha = .80 for authority ranking, and Cronbach's Alpha = .67 for market pricing (each 4 items). We conducted a confirmatory factor analysis using the R package "lavaan" (R version 4.1.2.; Rosseel et al., 2012). Model fit statistics indicated a satisfactory fit: $\chi^2(87)$ = 232, p < .001; CFI = .93, RMSEA = .07, SRMR = .08.

The average variance extracted is .54 (factor 1), .58 (factor 2), and .43 (factor 3). Values for composite reliability range from .75 to .91, which exceeds .70, indicating that all the items consistently measure their corresponding constructs. To assess discriminant validity, we furthermore tested the Fornell-Larcker Criterion, where all conditions are satisfied, suggesting an acceptable construct validity.

The correlations between the human-AI relationship are as in study 1: market pricing is positively correlated with peer bonding (r = .41, N = 362, p < .001) and authority ranking (r = .42, N = 362, p < .001). No significant correlation was found between authority ranking and peer bonding (r = −.07, N = 362, p = .156).

Testing the Impact of the Naming and the Relation to Demographic Variables

To test whether the naming of the conversational AI influenced human-AI relationship perception (H4) and whether the hierarchy of the relationship perceptions replicated study 1 (H1) we performed a 2 (naming–between) x 3 (relationship type–within) mixed-ANOVA and simple comparison of the naming factor**.** The results with Huynh-Feldt (HF) correction showed a significant main effect of the relationship type $F(1.61,578.43)$ = 383.3,3 p < .001, η_p^2 = .52). Participants saw their relationship with the conversational AI as more strongly characterized by authority ranking (M = 4.86, SD = 1.59) than by market pricing (M = 4.35, SD = 1.35) and more than by peer bonding (M = 2.42, SD = 1.33), all ps < .001. No naming by relationship mode interaction was found, $F(1.61,578.43)$ = 1.31, p = .267. Thus, the main effect of relationship type was not contingent on the naming.

Unexpectedly, independent of the relationship type, the naming of the AI affected the extent to which all three types were perceived, $F(1,360)$ = 4.40, p = .037, η_p^2 = .01 , indicating that mean values for all relationship modes were higher in the voice assistant condition.

To test whether the perceived relationship mode is related to demographic factors, we regressed each of the three scores separately on age, gender of user, household specifications (single vs with kids), level of education, employment (employed vs unemployed), technological knowledge, as well as the gender and screen settings of the conversational AI. The use of a screen with the smart speaker had a negative relationship with market pricing, $\beta = -.12$, $t(349) = -2.29$, $p = .023$. Furthermore, we found a negative relationship of age with authority ranking, $\beta = -.14$, $t(349) = -2.49$, $p = .013$, as well as a positive relationship with market pricing, $\beta = .11$, $t(349) = 2.01$, $p = .046$. Finally, we found a significant positive correlation of educational level with authority ranking, $\beta = .12$, $t(349) = 2.18$, $p = .030$, and a negative one with peer bonding, $\beta = -.14$, $t(349) = -2.68$, $p = .008$.

Discussion

Study 2 aimed at further testing the applicability of the RMT to human-AI relationships. The results of study 2 replicate the findings of study 1 regarding the dimensional structure and the hierarchy of the dimensions. The naming did not affect the hierarchy, not supporting H4. The main effect of naming occurred less relevant to us because the effect size was rather small and the effect occurred across all three relationship dimensions. Older and more educated people showed less peer bonding and market pricing but more authority ranking. Otherwise, the relationship perceptions were unrelated to the assessed demographic factors. In a nutshell, the relationship perceptions were largely unmoderated, validating that our approach is well suited to describe the human-AI relationships across different contexts.

General Discussion

How do humans perceive their relationship to conversational AI? By applying Fiske's (1992) multidimensional relational models theory (RMT), we aim to investigate human-AI relationships in relation to system perception and user characteristics.

First, the factor structure of our human-AI relationship questionnaire indicated that the relational modes Fiske (1992) suggests for human-human relationships can be applied to human-AI relationships, with one restriction: The more emotional modes of communal sharing and equality matching which characterize closer human relationships cannot be differentiated in the perception of the relationship between users and conversational AI. In both studies reported here, they converged on one factor that we called *peer bonding*. Overall, users perceived their relationships with conversational AI as predominantly characterized by authority ranking and market pricing. Peer bonding was much weaker than the two rational relationship modes mentioned before. These results indicate that users may perceive the human-AI relationship on similar dimensions as relationships to humans; however, they were less differentiated and less characterized by the emotional modes, which Fiske called communal sharing and equality matching.

These perceptions seem to be largely independent of the label used for the device and most demographic and system features, such as the gender of the conversational AI or household characteristics. This stresses that the relationship dimensions can be broadly applied and are relatively stable. At the same time, the only predictors at the user and system level that we could identify are age and education. Future research should explore psychological concepts such as personality, loneliness, etc., to identify individual differences in predicting human-AI relationship perception.

Furthermore, based on the broad availability of large language models such as ChatGPT[4] by OpenAI, conversational agents might become more performant, and conversational AI will likely appear more human-like. This opens the door to stronger human-AI relationships. Research monitoring these developments should also consider the multidimensional structure of perceived human-AI relationships.

As previously explained, many earlier studies in the broader field of human-machine interaction assumed the rational versus emotional dichotomy, for instance, warmth versus competence or companion versus assistant (Glikson & Woolley, 2020; Malle & Ullman, 2021). This matches the distinction between peer bonding and authority ranking. However, we found that market pricing, which denotes a rational and equal (other than hierarchical) dimension, enriches the current understanding, dominated by single or two-dimensional approaches, by adding another dimension.

This may be an interesting finding for developers of voice user interfaces. Although authority ranking and market pricing dimensions seem to have much in common at first (for instance, characterized as rational), they differ in one design-relevant important aspect: namely, the agency or responsibility people attribute to the conversational AI system. This likely impacts what tasks users use the conversational AI for. Does a request necessitate less machine intelligence, such as turning on a light, or does a request necessitate more machine intelligence, for instance, when shopping for goods with a conversational AI? We can imagine that for more sophisticated interactions, such as shopping, which involves multiturn dialogues, designers need to understand how they can trigger users to assume some reciprocal responsibility in interacting with the conversational AI.

We will elaborate on three key findings from the correlational analyses: (1) Low predictive value of authority ranking despite its high prevalence, (2) the predictive value of market pricing for trust, inclusion of AI in self, competence concerns and affinity to technology, and (3), highest predictive value of peer bonding for system perception (except trust), despite its low prevalence.

First, the highest mean values for authority ranking imply that a majority of users perceived their relationship as rational, task-oriented, and hierarchical. Given that the relation to conversational AI was dominantly perceived as authority ranking, we were surprised to see no unique correlations with any system perception variables apart from the number of purposes. This absence renders authority ranking a not very informative dimension regarding system perception and user characteristics, which might be partly due to its high prevalence.

Second, market pricing is both rational and task-oriented but not as hierarchically perceived as authority ranking. According to RMT, individuals who perceive a relationship to be characterized by market pricing perform a cost-benefit analysis on what they invest (time, money, engagement, etc.) and what they get out of the relationship—here by using the smart device. Given the rational or calculative characteristics, we were surprised to see correlations of market pricing with both rational variables, such as competence concerns, and variables we initially classified as emotional or have emotional components, for instance, inclusion of AI in the self. However, in line with current criticism against affective accounts of trust in AI (i.e., the technology is reliable in the sense that it works well, see, for instance, Ryan, 2020), it is plausible that trust and inclusion of AI in self can be considered more rational than emotional. More specifically, inclusion of conversational AI in self could also be understood as including the conversational AI as a part of users' daily routines. Overall, market pricing seems to be a more informative dimension than authority ranking, especially regarding the rational accounts of system perception. The findings so far further suggest that users form a relationship with conversational AI that is partly hierarchical but also partly equal.

Finally, despite the fact that peer bonding had the lowest mean, it was the strongest predictor of psychological distance, anthropomorphism, and perceived warmth—all variables that are of critical importance in studies of human-machine interaction, as laid down in the theoretical background. Thus, peer bonding seems to be the most informative dimension regarding system perception, especially emotional accounts thereof. Notably, there may be differences in the perception of this particular dimension between different off-the-shelf conversational AI. For instance, Kuzminykh et al. (2020) have shown in a qualitative study that Alexa is perceived differently—more friendly and warm—than Siri or the Google Assistant. It might be possible that people differ in their perception due to the different naming or gender settings (Abercombie et al., 2021). Although we could not identify such influence in study 2, we recommend considering this in future research.

Strengths and Limitations for Research and Practice

Unlike other studies, we did not focus on one specific role or the role intended by the designer (e.g., companion bots or (para-)friendships). Instead, we use a comprehensive model of human-to-human relationships applied to off-the-shelf conversational AI. These are arguably designed with multiple, ambiguous roles and can perform an increasing variety of tasks with increasing social dynamics, such as multiturn dialogues in voice shopping. Thus, even merely transactional interactions, like turning on the lights, are to some degree relational as we describe it with authority ranking.

Our study contributes in at least three ways: First, we discuss theories about relational approaches to studying human-AI interaction, expanding and integrating knowledge. Specifically, we are expanding the current literature by enriching the traditional rational-emotional or rational-transactional approaches (e.g., Xu & Li, 2022) of studying human-AI relationship perception with another dimension, which puts market pricing and peer bonding, next to authority ranking, into the spotlight. Secondly, we are expanding the methodological toolkit to study relationship modes quantitatively with the human-AI relationship questionnaire adapted from Haslam and Fiske (1999).

Thirdly, we are exploring how the relational models fit into the frameworks in the broader field of human-machine interaction, providing further empirical evidence of their relevance.

This approach can potentially offer practitioners insights from a design-based perspective. It mirrors the increasing social dynamics between humans and conversational AI and may be more efficient for designing toward the variety of tasks specifically off-the-shelf conversational AI can be used for. For example, what relational modes are more relevant for request and return dialogues, such as setting a reminder, versus interactions characterized by multiturn conversations in a specific moment, e.g., shopping or asking for music recommendations for a particular occasion? We propose a more functional approach, i.e., relating the design (e.g., language style) to the different tasks. However, this research approach is still in its infancy. We urge future research to adopt this approach and gather further empirical evidence to make conclusive claims for (design-) practice.

Several limitations should be taken into account. First, this cross-sectional study is investigating concepts, such as trust or relationship perception, that are likely to be time-sensitive (Glikson & Wooley, 2020; Seymore & van Kleek, 2021). Further studies should consider a longitudinal design to explore the effects over time and test for causality. Second, our sample was mainly from the UK, providing little demographic variation. Future research should complement these findings with a sample composed of various ethnic backgrounds.

Finally, compared to Langer et al. (2022), who found that the terminology of automated decision-making systems affects human perception, we could not identify a meaningful influence of the naming of the conversational AI, same for the gender settings. However, we believe that future studies should carefully consider both issues. In addition, it is foreseeable that more users will be able to change voice or gender settings, potentially influencing relationship perception.

Furthermore, future studies can use the current framework to investigate the impact on user behavior, which should also include behavioral data, such as audio scripts, to analyze actual behavior rather than relying on self-report measures only (e.g., Gao et al., 2018).

Conclusion

As conversational AI in fact *is* a tool, AI is not *just* a tool. The human-AI relationship was to a great extent perceived as a rational command-and-execution mode (authority ranking), and—in line with the suggestions in the introduction—users also displayed a less pronounced tendency to perceive their relationship to the conversational AI as emotional or peer-like (peer bonding). However, we found it most interesting that another prevalent dimension of human-AI relationships was perceived, which can be described as more rational and egalitarian. Highlighting these elements of exchange that characterize market pricing suggests that the conversational AI was ascribed or granted some agency or hierarchical equivalence. Consequently, authority ranking takes a back seat when investigating user perceptions and user characteristics. Borrowing from technical jargon, it looks like authority ranking is the default setting—hence, “servant by default”—while peer bonding and market pricing are the advanced or custom settings that users configure along the way.

Conflict of Interest

The authors have no conflicts of interest to declare.

Authors’ Contribution

Marisa Tschopp: conceptualization, data curation, formal analysis, investigation, writing—original draft. **Miriam Gieselmann:** formal analysis, writing—review & editing. **Kai Sassenberg:** conceptualization, data curation, formal analysis, supervision, writing—review & editing.

Footnotes

[1] With conversational artificial intelligence (conversational AI) we refer to AI systems users can talk to with voice. They are also often called smart personal assistants, virtual/digital assistants, or voice assistants (Kulkarni et al., 2019).

[2] Of note, we derived the predictions after the data was collected. In this sense, the current study should be considered exploratory.

[3] We thank the anonymous reviewer for suggesting this prediction.

[4] https://openai.com/

References

Abercrombie, G., Cercas Curry, A., Pandya, M., & Rieser, V. (2021). Alexa, Google, Siri: What are your pronouns? Gender and anthropomorphism in the design and perception of conversational assistants. In M. Costajussa, H. Gonen, C. Hardmeier, & K. Webster (Eds.), *Proceedings of the 3rd workshop on gender bias in natural language processing* (pp. 24–33). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.gebnlp-1.4

Amazon. (2021, February 2021). *Customers in India say "I love you" to Alexa 19,000 times a day*. Amazon. https://www.aboutamazon.in/news/devices/customers-in-india-say-i-love-you-to-alexa-19-000-times-a-day

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In Y. Nakano, M. Neff, A. Paiva, & M. Walker (Eds.), *Intelligent virtual agents* (pp. 126–138). Springer. https://doi.org/10.1007/978-3-642-33197-8_13

Bowlby, J. (1979). The Bowlby-Ainsworth attachment theory. *Behavioral and Brain Sciences, 2*(4), 637–638. https://doi.org/10.1017/S0140525X00064955

Buschke, D. (2021, July 6). *So erklären deutsche Nutzer ihre Zuneigung für Alexa – Alexa ich liebe dich.* [This is how German users declare their affection for Alexa - Alexa I love you] Smarthome Assistant. https://www.smarthomeassistent.de/so-erklaeren-deutsche-nutzer-ihre-zuneigung-fuer-alexa-alexa-ich-liebe-dich/

Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships*, *38*(1), 279–300. https://doi.org/10.1177/0265407520959463

Dautenhahn, K. (2007). Methodology & themes of human-robot interaction: A growing research field. *International Journal of Advanced Robotic Systems*, *4*(1), 103–108. https://doi.org/10.5772/5702

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, *99*(4), 689–723. https://doi.org/10.1037/0033-295X.99.4.689

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction*, *35*(6), 456–467. https://doi.org/10.1080/10447318.2018.1456150

Funk, G., Hofstetter, R., Kälin, F., Jörg, D., Lopetrone, R., & Bruetsch, C. (2021). Voice First Barometer Schweiz 2020. 1–48. https://www.unilu.ch/fakultaeten/wf/institute-zentren-forschungsstellen/institute-of-marketing-and-analytics/swiss-consumer-studies/voice-barometer-studie-2020/

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication*, *1*, 71–86. https://doi.org/10.30658/hmc.1.5

Gao, Y., Pan, Z., Wang, H., & Chen, G. (2018). Alexa, my love: Analyzing reviews of Amazon Echo. In *2018 IEEE SmartWorld, ubiquitous intelligence & computing, advanced & trusted computing, scalable computing & communications, cloud & big data computing, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (pp. 372–380). IEEE. https://doi.org/10.1109/SmartWorld.2018.00094

Gilad, Z., Amir, O., & Levontin, L. (2021). The effects of warmth and competence perceptions on users' choice of an AI system. In Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, & S. Drucker (Eds.), *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. https://doi.org/10.1145/3411764.3446863

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, *24*(4), 1494–1509. https://doi.org/10.1016/j.chb.2007.05.007

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior, 90*, 343–350. https://doi.org/10.1016/j.chb.2018.08.009

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society*, *22*(1), 70–86. https://doi.org/10.1177/1461444819858691

Han, S., & Yang, H. (2018). Understanding adoption of intelligent personal assistants. *Industrial Management & Data Systems*, *118*(3), 618–636. https://doi.org/10.1108/IMDS-05-2017-0214

Haslam, N., & Fiske, A. P. (1999). Relational models theory: A confirmatory factor analysis. *Personal Relationships, 6*(2), 241–250. https://doi.org/10.1111/j.1475-6811.1999.tb00190.x

Hepp, A. (2020). Artificial companions, social bots and work bots: Communicative robots as research objects of media and communication studies. *Media, Culture & Society, 42*(7–8), 1410–1426. https://doi.org/10.1177/0163443720916412

Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, *19*(3), 215–229. https://doi.org/10.1080/00332747.1956.11023049

Hu, P., Lu, Y., & Wang, B. (2022). Experiencing power over AI: The fit effect of perceived power and desire for power on consumers' choice for voice shopping. *Computers in Human Behavior, 128*, Article 107091. https://doi.org/10.1016/j.chb.2021.107091

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

Ki, C.-W., Cho, E., & Lee, J.-E. (2020). Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Computers in Human Behavior*, *111*, Article 106412. https://doi.org/10.1016/j.chb.2020.106412

Kim, A., Cho, M., Ahn, J., & Sung, Y. (2019). Effects of gender and relationship type on the response to artificial intelligence. *Cyberpsychology, Behavior, and Social Networking*, *22*(4), 249–253. https://doi.org/10.1089/cyber.2018.0581

Knapp, M. L. (1978). *Social intercourse: From greeting to good-bye.* Allyn and Bacon.

Kulkarni, P., Mahabaleshwarkar, A., Kulkarni, M., Sirsikar, N., & Gadgil, K. (2019). Conversational AI: An overview of methodologies, applications & future scope. In *2019 5th International conference on computing, communication, control and automation (ICCUBEA)* (pp. 1–7). IEEE. https://doi.org/10.1109/ICCUBEA47591.2019.9129347

Kuzminykh, A., Sun, J., Govindaraju, N., Avery, J., & Lank, E. (2020). Genie in the bottle. In R. Bernhaupt, F. F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. https://doi.org/10.1145/3313831.3376665

Langer, M., Hunsicker, T., Feldkamp, T., König, C. J., & Grgić-Hlača, N. (2022). "Look! It's a computer program! It's an algorithm! It's AI!": Does terminology affect human perceptions and evaluations of algorithmic decision-making systems? In S. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. Drucker, J. Williamson, & K. Yatani (Eds.), *CHI Conference on Human Factors in Computing Systems* (pp. 1–28). ACM. https://doi.org/10.1145/3491102.3517527

Li, X., & Sung, Y. (2021). Anthropomorphism brings us closer: The mediating role of psychological distance in user–AI assistant interactions. *Computers in Human Behavior*, *118*, Article 106680. https://doi.org/10.1016/j.chb.2021.106680

Lopatovska, I., & Williams, H. (2018). Personification of the Amazon Alexa. In C. Shah, N. J. Belkin, K. Byström, J. Huang, & F. Scholer (Eds.), *CHIIR '18: Proceedings of the 2018 Conference on Human Information Interaction & Retrieval* (pp. 265–268). ACM. https://doi.org/10.1145/3176349.3176868

Malle, B. F., & Ullman, D. (2021). A multidimensional conception and measure of human-robot trust. In Chang S. Nam & Joseph B. Lyons (Eds.), *Trust in human-robot interaction* (pp. 3–25). Elsevier. https://doi.org/10.1016/B978-0-12-819472-0.00001-0

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa … examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, *99*, 28–37. https://doi.org/10.1016/j.chb.2019.05.009

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers*. Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In C. Plaisant (Ed.), *CHI '94: Conference Companion on Human Factors in Computing Systems* (p. 204). ACM. https://doi.org/10.1145/259963.260288

Panetta, K. (2020, October 19). *Gartner top strategic technology trends for 2021*. Gartner. https://www.gartner.com/smarterwithgartner/gartner-top-strategic-technology-trends-for-2021

Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior, 140*(2), 107600. https://doi.org/10.1016/j.chb.2022.107600

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing*, *38*(4), 626–642. https://doi.org/10.1002/mar.21457

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). "Alexa is my new BFF". In G. Mark, S. Fussell, C. Lampe, m. c. schraefel, J. P. Hourcade, C. Appert, & D. Wigdor (Eds.), *CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2853–2859). ACM. https://doi.org/10.1145/3027063.3053246

Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, *109*, Article 106359. https://doi.org/10.1016/j.chb.2020.106359

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling*. Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Ryan, M. (2020). In AI we trust: Ethics, artificial intelligence, and reliability. *Science and Engineering Ethics, 26*(5), 2749–2767. https://doi.org/10.1007/s11948-020-00228-y

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity, 1*(4), 353–376. https://doi.org/10.1080/152988602760328012

Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management, 35*(7–8), 693–715. https://doi.org/10.1080/0267257X.2019.1596970

Seymour, W., & van Kleek, M. (2021). Exploring interactions between trust, anthropomorphism, and relationship development in voice assistants. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), Article 371. https://doi.org/10.1145/3479515

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion—A study of human-chatbot relationships. *International Journal of Human-Computer Studies*, *149*, Article 102601. https://doi.org/10.1016/j.ijhcs.2021.102601

Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S., Jung, E. H., Waddell, T. F., & Kim, K. J. (2017). Cheery companions or serious assistants? Role and demeanor congruity as predictors of robot attraction and use intentions among senior citizens. *International Journal of Human-Computer Studies*, *97*, 88–97. https://doi.org/10.1016/j.ijhcs.2016.08.006

Taylor, D., & Altman, I. (1987). Communication in interpersonal relationships: Social penetration processes. In M. E. Roloff & G. R. Miller (Eds.), *Interpersonal processes: New directions in communication research* (pp. 257–277). Sage Publications.

The Guardian. (2020, July 20). *'Alexa, I love you': How lockdown made men lust after their Amazon Echo*. https://www.theguardian.com/technology/2020/jul/20/alexa-i-love-you-how-lockdown-made-men-lust-after-their-amazon-echo

van Berkel, N., Skov, M. B., & Kjeldskov, J. (2021). Human-AI interaction. *Interactions*, *28*(6), 67–71. https://doi.org/10.1145/3486941

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social cognition unbound: Insights into anthropomorphism and dehumanization. *Current Directions in Psychological Science*, *19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Xie, T., & Pentina, I. (2022). Attachment theory as a framework to understand relationships with social chatbots: A case study of Replika. In T. Bui (Ed.), *Proceedings of the 55th Hawaii International Conference on System Sciences*. ScholarSpace. https://doi.org/10.24251/HICSS.2022.258

Xu, S., & Li, W. (2022). A tool or a social being? A dynamic longitudinal investigation of functional use and relational use of AI voice assistants. *New Media & Society.* Advance online article. https://doi.org/10.1177/14614448221108112

Appendix

Human-AI Relationship (Haslam & Fiske, 1999)

1. There is a moral obligation to act kindly to each other
2. Decisions are made together
3. You tend to develop similar attitudes and behaviors
4. It seems you have something unique in common
5. The two of you belong together
6. Some requests are granted in anticipation of something in return
7. "One-Person, one vote"' is the principle for making decisions
8. You take turns doing what the other wants.
9. You are like peers or fellow co-partners
10. One of us is entitled to more than the other
11. One directs the work, the other pretty much follows
12. You are like leader and follower
13. One is above the other in a kind of hierarchy
14. What you get is directly proportional to how much you give
15. You have a right to a fair rate of return for what you put into this interaction
16. You expect the same return on your investment other people get
17. Your interaction is a strictly rational cost-benefit analysis

Trust in Conversational AI (cAI; Jian et al., 2000)

1. The cAI is deceptive
2. [...] behaves in an underhanded manner
3. I am suspicious of [...]'s intent, action, or output
4. I am wary of [...]
5. [...] 's action will have a harmful or injurious outcome
6. I am confident in [...]
7. [...] provides security
8. [...] has integrity
9. [...] is dependable
10. [...] is reliable
11. I can trust [...]
12. I am familiar with [...]

Anthropomorphism (Waytz et al., 2010)

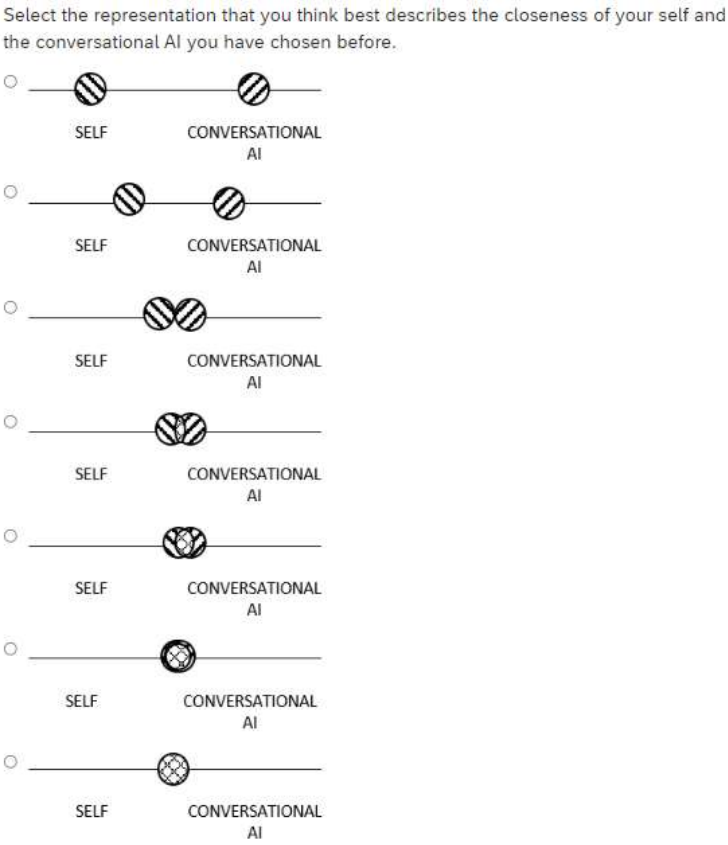
1. To what extent does the cAI have thoughts of its own?
2. To what extent does [...] have intentions?
3. To what extent does [...] have a free will?
4. To what extent does [...] have a consciousness?
5. To what extent does [...] have desires?
6. To what extent does [...] have values and norms?
7. To what extent does [...] experience emotions?

Psychological Distance (Li & Sung, 2021)

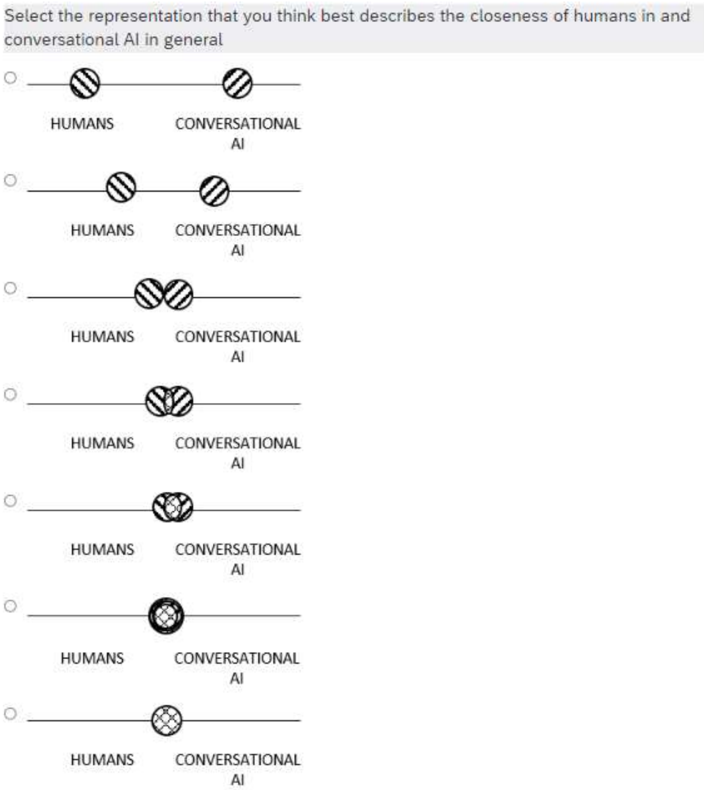
1. [...] is similar to me
2. [...] is psychologically close to me

Inclusion of AI in Self (Schubert & Otten, 2002)

1. Tick the representation that you think best describes the closeness of humans and digital assistants



2. Tick the representation that best describes your own closeness to the digital assistant



Affinity to Technology (Franke et al., 2019)

1. I like to occupy myself in greater detail with technical systems.
2. I like testing the functions of new technical systems.
3. I predominantly deal with technical systems because I have to.
4. When I have a new technical system in front of me, I try it out intensively
5. I enjoy spending time becoming acquainted with a new technical system.
6. It is enough for me that a technical system works; I don't care how or why.
7. I try to understand how a technical system exactly works.
8. It is enough for me to know the basic functions of a technical system.
9. I try to make full use of the capabilities of a technical system

Perceived Competence (Pitardi & Marriot, 2021)

1. I think the cAI is effective
2. I think [...] is intelligent
3. I think [...] is competent

Perceived Warmth (Pitardi & Marriot, 2021)

1. I think [...] is helpful
2. I think [...] is warm
3. I think [...] has good intentions

Frequency Usage (Funk et al., 2021)

1. How often do you use conversational AI?

Experience of Use (Funk et al., 2021)

1. Since when do you use conversational AI?

Purpose of Usage (Funk et al., 2021)

1. To retreive information (e.g., How is the weather tomorrow?)
2. To navigate (e.g., How long do I have from home to work by bike?)
3. To locate (e.g., Where is the next Italian restaurant?)
4. To control (e.g., Call my sister!)
5. To shop (e.g., Can you buy me flour?)
6. To entertain (e.g., Tell me a joke!)
7. Other purposes:

Barriers of Usage (Funk et al., 2021)

1. I don't want "someone" listening in all the time
2. I am concerned about privacy
3. I see no advantage in it
4. I find it uncomfortable/awkward to talk to a device
5. The conversational AI is often inaccurate in its statements
6. The conversational AI often does not understand me
7. The conversational AI cannot do what I expected it to do
8. Are there any other reasons that keep you away from using conversational AI more often or at all?

Demographic Variables

1. Please indicate your age
2. Please indicate your gender
3. Please Indicate your English language skills
4. In which country do you currently reside
5. What device have you used to answer this questionnaire

About Authors

Marisa Tschopp is a corporate researcher at scip AG (Zurich, Switzerland) and PhD student in the Social Processes lab at the IWM (Leibniz-Institut für Wissensmedien, Tübingen, Germany). Her research investigates users' perception of human-like characteristics of conversational artificial intelligence (AI), e.g., the human-AI relationship, trust in AI, and the influence on consumer behavior, such as voice shopping.

https://orcid.org/0000-0001-5221-5327

Miriam Gieselmann is a researcher and PhD student in the Social Processes lab at the IWM (Leibniz-Institut für Wissensmedien, Tübingen, Germany. Her research focuses on the question how users' subjective perceptions of human-like characteristics in Artificial Intelligence (AI)-based technical systems influence the acceptance of and interaction with these technologies.

https://orcid.org/0000-0002-4245-9237

Kai Sassenberg is head of the Social Processes lab at the IWM and full professor at the University of Tübingen (Germany). His research focuses on emotion, motivation, and self-regulation in the context of (a) social power and leadership, (b) the processing of (in-)correct information (e.g., conspiracy theories), and (c) the change of strong attitudes such as stereotypes and norms. In addition, he is interested in the acceptance of digital technologies – for instance artificial intelligence.

https://orcid.org/0000-0001-6579-8250

Correspondence to
Marisa Tschopp, scip AG, Badenerstrasse 623, 8048 Zürich, Switzerland, mats@scip.ch