Designing LLM-Based Conversational Agents for Prosocial Decisions: The Roles of Warmth and Competence

Completed Research Paper

Pascal Heßler
Karlsruhe Institute of Technology
Kaiserstraße 93, 76133 Karlsruhe
Germany
Pascal.Hessler@kit.edu

Tobias Blay
University of Goettingen
Platz der Göttinger Sieben 5, 37073
Göttingen Germany
Tobias.Blay@uni-goettingen.de

Jella Pfeiffer
Karlsruhe Institute of Technology
Kaiserstraße 93, 76133 Karlsruhe
Germany
Jella.Pfeiffer@kit.edu

Sebastian Hafenbrädl
IESE Business School
Avenida Pearson 21, 08034 Barcelona
Spain
SHafenbraedl@iese.edu

Abstract

As conversational agents increasingly support morally consequential decisions, it remains unclear whether anthropomorphic design promotes prosocial behavior. Integrating self-humanization theory with the stereotype content model, we examine how anthropomorphic cues shape perceived warmth, perceived competence and, in turn, algorithm aversion and prosocial behavior. In a preregistered, incentivized between-subjects experiment, 521 participants interacted with either a human-like conversational agent, a computer-like conversational agent, or a static text-based control before allocating money to an animal-related charity. Anthropomorphic cues increased both perceived warmth and perceived competence. However, only perceived competence was significantly associated with lower algorithm aversion and higher donation amounts, though the two paths did not differ significantly. Exploratory mediation analyses indicated an indirect association of anthropomorphic design with donations through competence. These findings qualify the presumed primacy of warmth in prosocial contexts and identify perceived competence as a pathway through which anthropomorphic conversational agents may support prosocial behavior.

Keywords: Conversational Agent, Warmth and Competence, Anthropomorphism, Prosocial Behavior

Introduction

Large language models (LLMs) are rapidly diffusing into everyday applications, heightening the need to understand how LLM-powered conversational agents (CAs) shape user behavior. In this study, the focal technological artifact is the LLM-based CA, whereas the LLM (i.e., generative artificial intelligence; AI) refers to the underlying model that generates the agent's responses. Despite growing use across domains such as healthcare (e.g., Q. Li et al., 2023) and finance (e.g., Huang et al., 2024), many users remain reluctant to rely on algorithmic advice – a phenomenon known as algorithm aversion (Dietvorst et al., 2015). A common root of algorithm aversion is the perception of algorithms as cold, calculating, and lacking

empathy (Liao & Huang, 2024). Such characteristics, referring to a perceived emotional detachment and analytical nature, can be advantageous for objective or computation-based tasks (Castelo et al., 2019; Seeger et al., 2021) and may even foster algorithm appreciation (Logg, 2017). However, in subjective or human-centered tasks, these same characteristics can backfire and increase aversion (Bigman & Gray, 2018).

The challenge is particularly acute in prosocial contexts – defined as decisions that benefit others (e.g., charitable giving) – where trust, empathy, and personal preferences are guiding principles (Eisenberg & Miller, 1987). Prosocial decisions are shaped by social preferences, moral concerns, and subjective judgments about how to benefit others (Berman et al., 2018). Thus, prior literature shows that individuals may favor decision processes that feel distinctly "human", allowing them to rely on human qualities (e.g., empathy) rather than purely algorithmic input (Heßler et al., 2022). However, despite the growing relevance of AI-mediated interactions in these domains, a fundamental tension arises: while prosocial decisions call for human judgment, they are increasingly shaped by CAs whose capacity to convey "humanness" and influence such decisions remains uncertain.

One promising lever for decreasing algorithm aversion is the use of anthropomorphic cues, such as designing CAs with relational language, a human name, or a profile image (Feine et al., 2019). It can make the CA appear more anthropomorphic, referring to the extent to which users subsequently experience the CA as human-like. Thus, these anthropomorphic cues invite users to attribute human qualities to the agent (Epley et al., 2007; Krach et al., 2008). Social perception of these cues organizes such attributions along two fundamental dimensions: warmth (e.g., benevolence, kindness, good intentions) and competence (e.g., ability, intelligence, effectiveness) (Fiske et al., 2002, 2007). Although many contemporary CAs already exhibit baseline human-like features (e.g., using natural language), further incorporating or emphasizing social cues (e.g., a face or a name) can help bridge the perceived human–machine gap (Seymour et al., 2025).

Prior research on human-AI interaction has examined both competence-related evaluations, such as accuracy and performance, and social responses to anthropomorphic design cues, such as trust (Christoforakos et al., 2021; Seeger et al., 2021). However, fewer studies have examined how warmth and competence relate to consequential user behavior in prosocial decision contexts. This is problematic since in prosocial contexts decisions are typically guided by empathy and moral concern, qualities closely associated with humans (e.g., warmth) (El Hedhli et al., 2023; Kim et al., 2019). As a result, it remains insufficiently understood how the design of CAs (i.e., the design of 'human' CAs) is consequential for prosocial behavior by perceiving the CAs as warm and competent. Regarding the dimension warmth, the stereotype content model (SCM) holds that it typically dominates social evaluation (Gilad et al., 2021; Wojciszke & Abele, 2008); whether that primacy extends to evaluating and, crucially, to relying on artificial agents in moral contexts is an open question.

The main goal of this study is to use charitable giving as an example for prosocial behavior to experimentally investigate the consequences on algorithm aversion and prosocial behavior of anthropomorphic design cues in CAs. We integrate self-humanization theory (Haslam, 2006) with the stereotype content model (SCM; e.g., Ahn et al., 2022), to develop a mediation model of how anthropomorphic design cues of an CA shape perceived warmth and competence of the CAs, which are subsequently associated with prosocial behavior (i.e., donation amount) and algorithm aversion. We propose that anthropomorphic cues can influence both dimensions by shaping perceptions of the CA's benevolent intentions and its ability to provide effective, context-sensitive guidance. To test our theoretical model, we designed an incentive-compatible experiment in which participants interacted with the CAs, and in which they made decisions that brought about real monetary consequences.

Our study contributes to the literature in two ways. First, we provide a more comprehensive and nuanced model of how people evaluate CAs to explain how anthropomorphic design cues shape human–computer interaction. Prior IS research shows that combinations of verbal, identity, and nonverbal cues shape perceived anthropomorphism, engagement, trust, and transactional outcomes (Schanke et al., 2021; Schuetzler et al., 2020; Seeger et al., 2021). More recent work identifies warmth and competence as distinct pathways through which chatbot anthropomorphism affects service evaluations and purchasing (Y. Li et al., 2025). We extend this literature by integrating self-humanization theory with the stereotype content model to connect anthropomorphic design cues, attributed humanness, warmth, and competence within a unified account of CA evaluation. This integration moves beyond documenting whether anthropomorphic

cues are beneficial by specifying how they activate distinct social judgments that shape algorithm aversion and behavior.

Second, we contribute by extending research into a prosocial context characterized by morally consequential decisions and, in doing so, refine theoretical assumptions about the role of warmth and competence. Prior literature suggests that prosocial behavior is primarily driven by affective responses linked to warmth, such as empathy and concern for others (Zhou et al., 2019). However, our findings qualify this assumption by showing a distinct pattern: only competence was significantly associated with both lower algorithm aversion and higher prosocial behavior in this context. This suggests that, in AI-mediated prosocial decision-making, individuals may place weight on the perceived ability of the agent to effectively translate intentions into meaningful outcomes, rather than on its perceived benevolence alone. Therefore, we highlight an important boundary condition of SCM, indicating that the importance of warmth and competence may shift in technology-mediated settings and, in particular, in contexts where decisions are enacted through CAs.

Theory

Our theoretical starting point is anthropomorphism: the process through which human-defining qualities are attributed to non-human objects, including CAs (Epley et al., 2007). CAs invite such attributions by design. Social cues (e.g.: verbal style, human identity markers such as a name and face, and nonverbal signals (Feine et al., 2019; Seeger et al., 2021), lead users to infer emotional capacity, intentionality, and agency, thereby perceiving the agent as more human-like, even when they know they are interacting with software (Brendel et al., 2023; Krach et al., 2008). What, exactly, can be attributed is clarified by self-humanization theory which distinguishes two core dimensions of humanness: human nature (e.g., emotionality, warmth) and human uniqueness (e.g., rationality, competence; Haslam, 2006). These dimensions capture the essential qualities that differentiate humans from non-human entities and thus define the attributes that may be ascribed when perceiving humanness. In prosocial contexts, such humanness attributions are particularly consequential, as they can activate social and moral responses typically reserved for human interaction partners (Banks, 2019).

To explain how these humanness attributions translate into social evaluations, we draw on the SCM, which posits that social perception is organized along two fundamental dimensions: warmth and competence (Cuddy et al., 2008; Fiske et al., 2002, 2007; Khadpe et al., 2020). These dimensions support rapid friend–foe assessments – whether others intend good or harm (warmth) and whether they can enact those intentions (competence; Fiske et al., 2007). The dimensions of self-humanization theory are conceptually related to, but not equivalent to, the SCM dimensions of warmth and competence. Warmth captures intention (e.g., kindness and benevolence); competence captures capability (e.g., skills and effectiveness). In the context of prosocial decision-making, these dimensions are especially relevant, as warmth shapes affective responses such as empathy and moral concern, while competence informs expectations regarding the effectiveness and impact of helping behavior. Moreover, people may apply social rules to CAs, allowing SCM to extend to the prediction and explanation of people's evaluations of CAs (Khadpe et al., 2020; Reeves & Nass, 1996). Research highlights that warmth dominates first impressions – *primacy of warmth* – because detecting intentions is evolutionarily more urgent and cognitively easier than judging the ability to enact those intentions (Wojciszke & Abele, 2008). Empirically, warmth explains more variance in global evaluations than competence (Wojciszke et al., 1998). Thus, the SCM posits that warmth and competence shape expectations, with warmth having a particular relevance (e.g., Kim et al., 2019).

Considered together, these perspectives offer a framework that links the design of CAs to perceived humanness and, subsequently, to the formation of socially meaningful judgments (e.g., donation to a charity): anthropomorphic cues, explains why CAs come to be treated as social actors, self-humanization theory clarifies which dimensions of humanness are invoked in this process, and the SCM provides a structured account of how these inferred human qualities are organized into fundamental social evaluations. The framework also positions our study relative to the closest prior work: Heßler et al. (2022) showed, for prosocial decisions, that attributions of humanness underlie algorithm aversion. Our model adds the social-perceptual layer through which such attributions become evaluations: warmth and competence. Additionally, it traces those evaluations through to incentivized prosocial behavior.

Hypotheses

We developed our research model outlined in Figure 1.[1] We argue that anthropomorphic cues raise perceived warmth through humanness attribution: the cues invite users to ascribe human-nature qualities (e.g.: emotional capacity, empathy, benevolent intent) to the CA (Epley et al., 2007; Haslam, 2006), and warmth is precisely the social evaluation of such qualities (Fiske et al., 2007). Empirical work supports each link in this chain: anthropomorphic design cues increase perceptions of the agent as a social actor (Araujo, 2018), and anthropomorphizing agents has been found to increase their perceived warmth directly (Kim et al., 2019; Z. Wang et al., 2016). Warmth-related qualities matter in our setting: empathy and warmth are central to charitable giving (Eisenberg & Miller, 1987; Zhou et al., 2019), and humanness attributions shape whether users accept support in prosocial decisions (Heßler et al., 2022). Thus, we hypothesize:

H1a. A CA designed with richer anthropomorphic cues will be perceived as warmer.

In a similar fashion, anthropomorphized CAs may be perceived as more competent, although this relationship is less straightforward than for warmth. According to the SCM, competence reflects whether an actor is seen as capable and effective. Thus, a more human-like CA may appear more competent because users attribute intentionality, understanding, and task-related ability to it. A link supported for smartphone assistants (Cheng, 2022) and for advice-giving CAs, where social presence raises perceived competence (Schmid et al., 2022). However, self-humanization theory suggests a more nuanced view. Some uniquely human attributes, such as logic and rationality, can also be shared with machines, meaning that even non-anthropomorphized agents may be perceived as competent in analytical tasks (Haslam, 2006). Yet in prosocial contexts, competence may depend not only on rational calculation but also on empathy, moral sensitivity, and the integration of facts and feelings – attributes less readily associated with machines (Berman et al., 2018; Bigman & Gray, 2018; Liao & Huang, 2024) . While the primacy of warmth in anthropomorphized interactions suggests that the effect on competence may be weaker than the effect on warmth (El Hedhli et al., 2023), anthropomorphic cues can still increase perceived competence by making the agent appear more intentional, understanding, and capable of context-sensitive guidance.

At the same time, this expectation is not unequivocal: emphasizing human-like qualities could also make the CA appear less machine-like and therefore less strongly associated with analytical precision, objectivity, or computational expertise. Acknowledging these competing theoretical expectations, we expect the context-sensitive and socially embedded nature of prosocial decision-making to make the positive competence-enhancing effect of anthropomorphism more likely. Therefore, we hypothesize:

H1b. A CA designed with richer anthropomorphic cues will be perceived as more competent.

Next, we consider the consequences that perceived warmth and competence might have on algorithm aversion. Algorithm aversion is the central concept when it comes to understanding whether users rely on and accept AI-driven recommendations. More broadly, whether they are willing to delegate parts of a decision to an agentic system (Baird & Maruping, 2021). Jussupow et al. (2020) define it as a "biased assessment of an algorithm which manifests in negative behaviors and attitudes towards the algorithm compared to a human agent" (p. 4). As an umbrella term, it tends to arise in high-stakes decisions (Longoni et al., 2019), after algorithm errors (Dietvorst et al., 2015), or when decisions are seen as more subjective than objective (Castelo et al., 2019) – as in charitable giving, often viewed as subjective (Batson et al., 1999; Loewenstein & Small, 2007).

One source of algorithm aversion is a lack of perceived humanness in the agent. A mechanism Heßler et al. (2022) demonstrated for prosocial decisions: individuals evaluate whether an interaction partner possesses human-defining attributes, and in the absence of such attributes, the agent is perceived as fundamentally "other". Building on the SCM, warmth and competence capture whether the agent is seen as having benevolent intentions (i.e., warmth) and the ability to act effectively (i.e., competence). Thus, deficits in perceived warmth and competence should directly contribute to algorithm aversion.

Furthermore, the SCM suggests stereotypes serve as cognitive shortcuts: portraying an anthropomorphized CA as warm and competent can bridge the human–algorithm gap (Cervellon et al., 2019; Chu et al., 2016).

[1] Our preregistration names two further CA evaluations: social presence and shared reality. The research model focuses on the two SCM dimensions.

Applying a human-agent stereotype may prompt attributions of human-like rationality rather than rigid rule-following, reducing negative perceptions of human–machine differences. Increasing perceived warmth and competence through anthropomorphic cues should therefore reduce the perceived human–machine divide by signaling that the CA possesses both benevolent intentions and the capacity for effective action. As a result, the agent is evaluated less as a purely mechanical system and more as a socially relevant actor. Both dimensions reflect core aspects of perceived humanness and should thus contribute independently to reducing algorithm aversion, leading to the following hypotheses:

H2a. Higher perceived warmth of the CA will decrease algorithm aversion.

H2b. Higher perceived competence of the CA will decrease algorithm aversion.

Furthermore, we expect higher levels of warmth and competence to translate into higher prosocial behavior (i.e., greater donation amount). This also links to a growing interest in IS research on technology-mediated prosocial decisions, from data donations (Alashoor et al., 2025) to charitable crowdfunding (Blicher et al., 2026). When a CA is perceived as warmth and more competent, users experience more engaging interactions, which can deepen their involvement with the charity decision and increase confidence in their choice. Warmth signals good intentions and alignment with the prosocial purpose of the decision, which may make the CA's guidance appear more socially appropriate. Competence signals that the CA can understand the user's preferences, evaluate the available charities, and provide effective guidance, which may increase willingness to act on its advice. Consistent with this logic, consumer research shows that people are more likely to buy from companies perceived as both warm and competent (Aaker et al., 2010). Given the inherent subjectivity of selecting a charity, users might particularly value a CA that can competently provide accurate, personalized, and appropriate guidance. Which in turn increases persuasiveness of its recommendation, which should positively influence the amount donated. Thus, perceived warmth and competence should not only shape evaluations but also directly translate into prosocial behavior. Therefore, we hypothesize:

H3a. Higher perceived warmth of the CA increases prosocial behavior.

H3b. Higher perceived competence of the CA increases prosocial behavior.

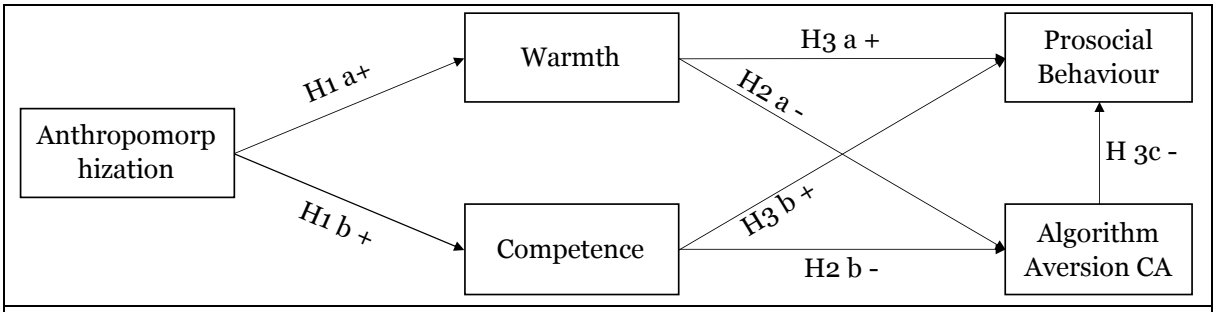


Figure 1. Research model

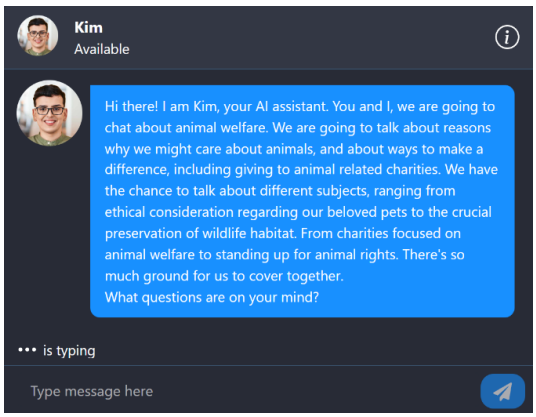
As argued, algorithm aversion reflects a negative tendency towards algorithms. From our theoretical perspective, it reflects a rejection of the CA as a socially meaningful, human-like actor. Aversion to a CA recommending charities can plausibly spill over to the charities it suggests. High algorithm aversion also fosters viewing the CA as a cold, calculating machine rather than a human-like entity. Such perceptions imply low attributed warmth and competence, thereby undermining the social evaluations that typically support prosocial behavior and that conflict with the values typically associated with charitable giving. Finally, people tend to seek social or parasocial relationships with advice sources (Prahl & van Swol, 2017); when algorithm aversion is high, forming such a relationship with the CA is harder, reinforcing negative perceptions of the system that are likely to be reflected in behavior. Thus, algorithm aversion represents a behavioral consequence of diminished perceived humanness, which in turn reduces willingness to donate. These considerations lead us to the following hypothesis:

H3c. Higher algorithm aversion towards the CA will decrease prosocial behavior.

Method

We manipulated CA anthropomorphization between-subjects through the three cue types of Seeger et al. (2021): verbal, human-identity, and nonverbal cues. Both CAs were implemented with ChatGPT (gpt-4o-2024-05-13) via a self-developed interface and differed only in system prompts, start messages, and interface cues (Figure 2):

- Human-like CA ("Kim"): prompts instructing high verbal anthropomorphism; a name and gender-neutral profile photo; a "typing…" indicator, available state.
- Computer-like CA ("Assistant System"): prompts instructing minimal verbal anthropomorphism; no name or image; a "generating…" indicator, online state.
- No-CA control: a single static text, presented instead of an interaction. We constructed it by extracting and synthesizing the central informational content of our own participants' dialogues from both CA conditions (using ChatGPT for extraction and drafting) and framed it as originating from an assistant system to hold the source constant.



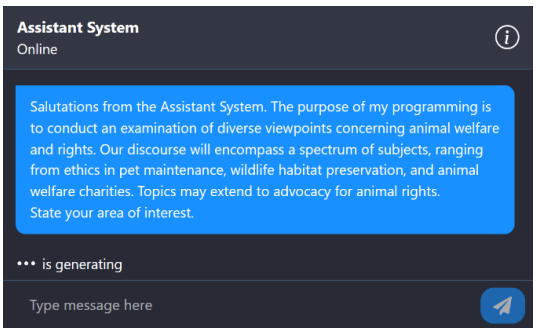


Figure 2. Start message and design of CA: left human-like, right computer-like

The manipulation is deliberately a bundle of cues: our aim is to compare a realistically anthropomorphized CA against a realistically machine-like one, not to isolate single cues.

A prestudy (N=174) verified the manipulation, prompted minor adjustments (e.g., emoticons did not increase anthropomorphization), found no uncanny-valley effect (MacDorman et al., 2009; Tinwell & Sloan, 2014), and indicated that animal-related charities engage participants most.

Procedure

The procedure of our experiments was structured as follows: (1) participants provided consent; (2) they completed an engagement task in which they viewed one of three images of endangered animals and wrote down their thoughts; (3) they received a briefing about the upcoming interaction with either a human or computer-like CA, or the static text in the No-CA condition, by an understanding check (Goodman et al., 2013); (4) participants in the CA conditions interacted with the agent for a minimum of eight message exchanges (threshold not disclosed), with conversation starters available before and during the interaction, while No-CA participants read the static text instead; (5) participants were introduced to the donation task, in which they allocated a budget of £18 between one of three animal-welfare charities (SICSA, PETA, or CROW) and keeping the money; (6) they selected a charity and decided on the allocation, with prosocial behavior operationalized as the donated amount. To make the decision incentive-compatible, a lottery gave participants a 10% chance that the budget would be paid out and allocated according to their choice (cf. DellaVigna et al., 2012) and (7) participants completed a questionnaire, including an attention check.

We obtained ethics approval from the IESE ethics committee and preregistered the study (https://aspredicted.org/KJ5_JD7). A pwrSEM power analysis A. Wang & Rhemtulla (2021) suggested N = 500, set the recruitment target at 600 to allow for exclusions.

All study materials, including measurement items and detailed analyses, are available from the authors upon request.

Participants

604 Prolific participants were recruited. The No-CA text had to be constructed from the CA interaction transcripts, so the third condition was fielded within four days of the others (including a weekend) rather than concurrently. We treat this as a limitation and report balance checks across treatments Table 1. After exclusion, the preregistered criteria (missing answers, n = 1; off-topic chats, n = 8; failed attention checks, n = 53) plus two disclosed data-integrity criteria not in the preregistration (multiple attempts, n = 13; incomplete log data, n = 8, the final sample consisted of 521 participants (51% female; age M = 41.8, SD = 13.2): human-like n = 173, computer-like n = 170, No-CA n = 178. The preregistered relative-speed rule was not applied as the original speed index was not included in the export by Sosci. Recomputing an approximate index on the original sample identifies 16 fast responders, and excluding all of them changed no conclusion.

Measures

All measures used a seven-point scale unless noted.

We check the manipulation on three complementary levels. (1) *Perceived anthropomorphization*: a simplified agency/experience scale based on Gray et al. (2007), because attributing mind is the core of anthropomorphism (Epley et al., 2007). (2) *Social presence* (Gefen & Straub, 2003), because anthropomorphic cues should make the interaction feel more like contact with a social actor, a perception distinct from mind attribution. (3) An objective language check: all CA messages scored with LIWC-22 on word count, affect, cognition, personal pronouns, and social processes. Verifying that the prompts changed what the CA actually wrote, independent of participants' perceptions.

The preregistration names four evaluations: *perceived warmth* (friendly, warm, kind, pleasant) and *perceived competence* (competent, effective, skilled, intelligent), four-item adaptations of Fiske et al. (2002) following Aragonés et al. (2015), *social presence* (above); and *shared reality* (8 items, Rossignac-Milon et al., 2021). The research model focuses on warmth and competence as the SCM dimensions.

The primary measure is a one choice-based item: preference between a human supporter and the experienced CA (1–7, higher = prefer the human), following the choice-based approach of Jussupow et al. (2020) and Longoni et al. (2019). We use the choice-based item as primary because algorithm aversion is defined *comparatively* (e.g.: preferring a human over an algorithm), and because of its proximity to behavior. In addition, six attitudinal items (Heßler et al., 2022) measured trust, appropriateness, and authenticity, each once toward a human helper and once toward the CA; these referent-specific items enter the measurement analyses in the Results. And finally p*rosocial behavior* was operationalized with the donation amount (£0–18).

In addition, we measured further outcomes and added control variables: Enjoyment (J. Moon & Kim, 2001), intention to reuse (Venkatesh et al., 2003) Affinity for Technology Interaction (ATI; Franke et al., 2019), Attitude towards AI (AtAI; Grassini, 2023), participant gender (binary, male = 1; from the panel's demographic export); Kim's perceived gender in the human-like arm.

Analytical Approach

We estimated all structural models with partial least squares-structural equation modelling (PLS-SEM) using the SmartPLS software (Ringle et al., 2024) with 5,000 bootstrap resamples throughout. Although the merits of PLS versus covariance-based SEM are debated (e.g., Goodhue et al., 2012; Hair et al., 2011), our goal is to validate our hypotheses rather than precise parameter magnitudes or global fit, making PLS-SEM especially suitable (Goodhue et al., 2012).

We used standard binary coding for the treatment conditions and controlled for individual differences by adding a path from AtAI to algorithm aversion and, given evidence that women spend more than men (De Wit & Bekkers, 2016), we added a path from gender to prosocial behavior (i.e., donation amount).

Results

Balance and Manipulation Check (Exploratory)

We checked age, sex, Prolific approval counts, affinity for technology, and prior attitude towards AI (all $p \geq 0.122$; Table 1), to confirm non-concurrent fielding of the No-CA condition compromised group equivalence.

Variable	No-CA	Computer-like	Human-like	Test	p
Age (years)	41.37 (13.41)	43.29 (12.25)	40.73 (13.71)	F(2, 518)=1.75	0.174
Attitude towards AI	5.06 (1.44)	5.20 (1.33)	5.14 (1.44)	F(2, 518)=0.40	0.669
Affinity for tech.	4.32 (1.17)	4.18 (1.21)	4.23 (1.16)	F(2, 518)=0.58	0.563
Prolific approvals	1426 (1551)	1828 (2047)	1680 (1928)	F(2, 518)=2.11	0.122
Male (%)	49.4%	54.1%	43.4%	chi2(2)=4.00	0.135

Table 1. Balance checks across the treatments

Note. Because the No-CA treatment was fielded non-concurrently, we compare all pre-treatment covariates across conditions; no test is significant at the 5% level.

Regarding manipulation checks, the human-like CA exceeded both other treatments on perceived anthropomorphization (1) (vs computer-like: diff = 0.73; vs No-CA: 1.00; both $p \leq 0.001$) and social presence (2) (0.73 and 0.69, $p < 0.001$), while the computer-like CA and No-CA did not differ significantly on either check ($p = 0.48$ and $p = 1.00$). Only the human-like CA significantly exceeded the others on these checks: natural-language interactivity alone did not significantly raise perceived anthropomorphism above static text. On the language level (3), utilizing the LIWC dictionary (Pennebaker et al., 2014; Tausczik & Pennebaker, 2010), we identified that the human-like CA produced more words (*diff* = 344.83; $t = 15.82$) showed higher scores on affect (*diff* = 1.88; $t = 10.10$), cognition (*diff* = 1.78; $t = 5.65$), personal pronouns (*diff* = 4.33; $t = 28.59$) and social processes (*diff* = 2.87; $t = 7.76$; *all* $p < 0.001$) compared to the computer-like CA.

Table 2 offers a descriptive overview of the main constructs.

Latent Construct	No-CA	Computer-like CA	Human-like CA
Perceived anthropomorphization	3.90 (1.76)	4.17 (1.81)	4.90 (1.80)
Social presence	4.04 (1.63)	4.01 (1.71)	4.73 (1.64)
Warmth	4.45 (1.52)	4.69 (1.61)	5.64 (1.25)
Competence	5.49 (1.26)	5.64 (1.32)	6.02 (1.14)
Algorithm aversion	4.52 (1.66)	3.93 (1.75)	4.02 (1.70)
Donation amount (in £)	7.49 (5.53)	7.67 (5.64)	8.72 (5.89)
Attitude towards AI (control)	5.06 (1.44)	5.20 (1.33)	5.14 (1.44)

Table 2. Descriptive statistics

Note: Cell entries are means (standard deviations). N = 521 (No-CA n = 178, computer-like n = 170, human-like n = 173).

Measurement Validation (Exploratory)

We assessed convergent and discriminant validity for all latent constructs. Cronbach's alphas and composite reliabilities exceeded the threshold of 0.70 (Hair et al., 2016), and average variance extracted values were above the minimum of 0.50 (Hair et al., 2016). Discriminant validity was supported by Fornell–Larcker criterion (Fornell & Larcker, 1981), and HTMT values below 0.85 (Henseler et al., 2015). In summary, our measures exhibited an adequate level of both convergent and discriminant validity.

Hypotheses Testing (Preregistered)

We estimated the research model (Figure 1), Figure 3 shows the empirical model with the estimates, and Table 3 summarizes the hypotheses.

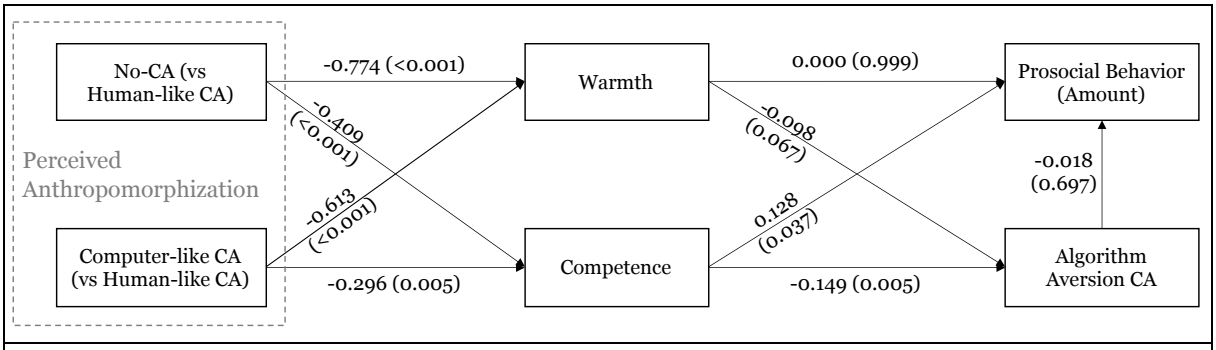


Figure 3. Empirical Model.

Note: Treatment paths are relative to human-like CA. Path-values are standardized; p-values presented in brackets are two-sided.

Relative to the human-like CA, both other conditions were perceived as less warm (No-CA: $\beta = -0.774$; computer-like: $\beta = -0.613$; both $p < 0.001$) and less competent ($\beta = -0.409$, $p < 0.001$; $\beta = -0.296$, $p = 0.005$), supporting H1a and H1b for the human-like contrasts. Meaning the human-like (i.e., more anthropomorphized) treatment increases perceived warmth and competence (the competence prediction was preregistered non-directionally). The computer-like condition did not differ from No-CA (warmth β = -0.161, $p = 0.144$; competence: β = -0.114, $p = 0.314$), consistent with the manipulation checks.

Regarding the hypotheses on algorithm aversion, competence has a significantly negative impact ($\beta = -0.149$, $p = 0.005$; H2b supported), whereas warmth fell short of the 5% level ($\beta = -0.098$, $p = 0.067$; (H2a not supported). The two paths do not differ significantly ($\chi^2(1) = 0.34$, $p = 0.558$).

In terms of prosocial behavior (i.e., donation amount), competence has a significantly positive effect (β = 0.128, $p = 0.037$; supporting H3b), while warmth ($\beta = 0.00$, $p = 0.999$; H3a not supported) and algorithm aversion (β = -0.018, $p = 0.697$; H3c not supported) are non-significant. The warmth and competence paths again do not differ significantly ($\chi^2(1) = 2.27$, $p = 0.132$).

Overall, prosocial behavior (i.e., donation amounts) shows unexpected patterns with respect to warmth and algorithm aversion, plausibly because they also reflect extraneous factors – participants' financial constraints, reluctance to spend, or dissatisfaction with the available charities.

Hypo.	Description of Hypotheses	β (p-value)	Valid?
H1a	No-CA vs Human-like -> Warmth	-0.774 (<0.001)	Yes
H1b	No-CA vs Human-like -> Competence	-0.409 (<0.001)	Yes
H1a	Computer-like vs Human-like -> Warmth	-0.613 (<0.001)	Yes
H1b	Computer-like vs Human-like -> Competence	-0.296 (0.005)	Yes
H2a	Warmth -> Aversion	-0.098 (0.067)	No
H2b	Competence -> Aversion	-0.149 (0.005)	Yes
H3a	Warmth -> Amount	-0.000 (0.999)	No
H3b	Competence -> Amount	-0.128 (0.037)	Yes
H3c	Aversion -> Amount	-0.018 (0.697)	No

Table 3. Overview hypotheses tests

Besides prosocial behavior (i.e., donation amount) and algorithm aversion, we collected two additional dependent variables: enjoyment and intention to reuse. An ANOVA post hoc test showed that the CAs did not differ significantly in enjoyment (*diff* = 0.27, $p < 0.42$), and both were more enjoyable than No-CA (human-like: *diff* = 0.81, $p < 0.001$; computer-like: *diff* = 0.54, $p = 0.01$). For intention to reuse, an ANOVA found only a marginal difference between No-CA and human-like at the 10% level (*diff* = 0.45, $p = 0.07$).

Robustness checks

As preregistered alongside PLS, we re-estimated the research model with CB-SEM, every substantive conclusion matches. Since the model fits poorly, we treat it as a preregistered convergence check on the path estimates, not as fit-based validation.

We also ran multiple exploratory robustness checks: Excluding controls (attitude towards AI and gender) lead to warmth significantly decreasing algorithm aversion, suggesting a support of H2a. Thus, attitude towards AI might serve as a boundary condition on the warmth path on algorithm aversion. Conclusions are unchanged when (a) replacing the experimental contrasts with perceived anthropomorphization (β = 0.78 on warmth, 0.625 on competence, both p < 0.001), (b) controlling for the CA's word count (computer-lie CA = 422.93; human-like CA = 767.76), which can act as potential confound (Y. Moon, 2000). We found no evidence that word count predicts warmth (β = *0.026,* p = 0.590) nor competence (β = *-0.041,* p = 0.459), and all paths remain stable. Summarizing the perception results are not explained by a simple linear association with word count. This is a descriptive adjustment: word count is treatment-induced, so conditioning on it cannot isolate a verbosity-purged anthropomorphism effect. Word count retains a direct association with donations (β = 0.0027, p = 0.021). Adding the preregistered ATI control (c) likewise changes nothing. And finally, addressing the manipulation check we estimated the model, contrasting the human-like condition against both others pooled (d), which leaves all conclusions unchanged.

A Complementary Measure of Algorithm Aversion (Exploratory)

As alternative to using the one choice-based item for algorithm aversion, the attitudinal aversion items measure trust, appropriateness, and authenticity toward the CA and toward a human helper, which permits a measurement analysis that sharpens the warmth result. Modelling the three CA-referenced items as the outcome (α = 0.86), both warmth (β = 0.138, p = 0.006) and competence (β = 0.338, p < 0.001) predicted more favorable evaluations of the CA. Modelling the human-minus-CA difference scores (α = 0.84), a comparative measure like our primary choice-based item, only competence remained (β = −0.201, p = 0.002; warmth: β = −0.036, p = 0.579). Warmth is thus associated with more favorable absolute evaluations of the CA, though competence's association on that measure is robustly larger (difference = −0.20, 95% CI [−0.37, −0.01], p = 0.038. The comparative preference for a human over the CA was significantly associated only with competence. The warmth–competence difference on the comparative measure is not itself significant.

Mediation Models (Exploratory)

Because the mediators were measured rather than manipulated, and were elicited in the post-task questionnaire, the indirect effects are interpreted as statistical associations consistent with the theorized pathways, not as causal mediation estimates.

To probe the volatility of prosocial behavior (i.e., donation amounts) and potential direct treatment effects, we ran supplemental analyses focusing on (a) prosocial behavior (donation amount) and (b) algorithm aversion. Starting with donation amount (a), an ANOVA across all three conditions showed no significant differences (F(2, 518) = 2.37, p = 0.094); the focused human-like vs computer-like comparison gave a difference of £1.05 (t(341) = 1.69; two-sided p = 0.092; one-sided p = 0.046, reported for the directional theoretical expectation), warranting further examination.

Given our theory that warmth and competence should affect prosocial behavior (i.e., donations) – and the weak total effect of the human-like condition on amount – we estimated a mediation model to clarify how the manipulation relates to donation amount. The mediation model and its results are depicted in Figure 4, again we include a path from gender to donation amount.

Bootstrap confidence intervals (5,000 resamples; bias-corrected) were used to assess indirect effects and mediations (Preacher & Hayes, 2004). The experimental condition had no significant direct effect (c_1 and c_2) on donation amount. In addition, we analyzed indirect effects, showing a significant effect through competence (a_2b_2 and a_4b_2), a pattern consistent with indirect-only mediation (Zhao et al., 2010), while warmth (a_1b_1 and a_3b_1) shows no effect. Furthermore, competence positively predicted prosocial behavior (i.e., donation amount) (b_2), whereas warmth did not (b1), consistent with the main model. Although the human-like CA increased warmth (a_1, a_3), warmth did not translate into higher prosocial behavior (i.e., donations; rejecting H3a), while competence did (supporting H3b).

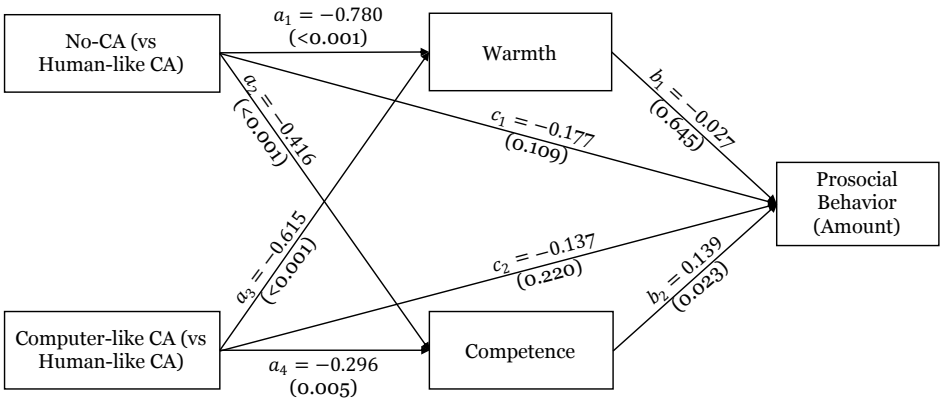


Figure 4. Mediation model prosocial behavior (i.e., donation amount)

As in the model before, the direct effect (c1 and c2) from the manipulation on algorithm aversion revealed no significant effects; interestingly, the paths have opposite signs. The negative sign of the c_2 path likely reflects statistical balancing: the CA variants have similar mean aversion, while the human-like CA reduces aversion indirectly via competence, yielding a compensating negative direct path and hinting at a possible omitted variable. As with prosocial behavior (i.e., donation amount), we identified an indirect effect through competence (a_2b_2 and a_4b_2), but not through warmth (a_1b_1 and a_3b_1), yielding an indirect-only pattern (Zhao et al., 2010).

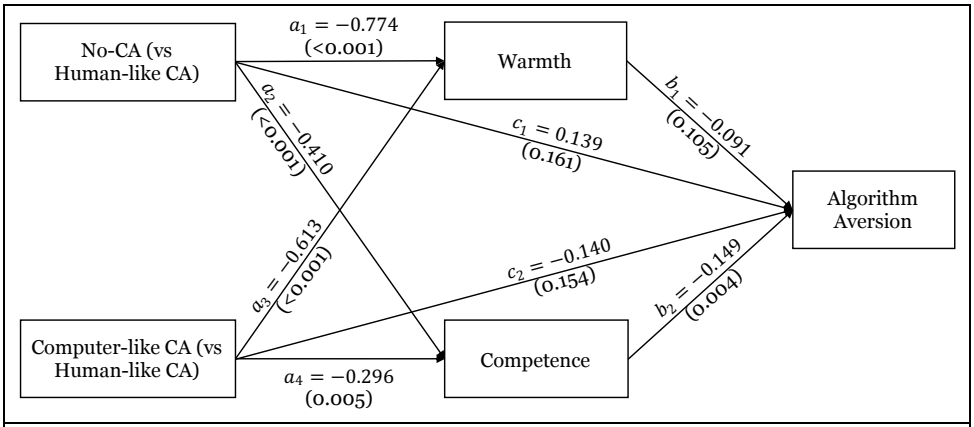


Figure 5. Mediation model algorithm aversion

In summary, across both outcomes, the additional models converge on a common mechanism: anthropomorphization does not exert a direct effect once controls are included, but it operates indirectly through perceived competence. Warmth increases under the human-like manipulation but does not significantly reduce algorithm aversion or increase prosocial behavior (i.e., donations) in these models, whereas competence consistently does. For algorithm aversion, both CA variants outperform the text-only control, but the human-like CA is not superior to the computer-like CA, the competence pathway again accounts for the observed reductions in aversion.

Discussion

Understanding the role of anthropomorphism is gaining relevance as CAs are becoming increasingly involved in morally consequential decisions. Integrating self-humanization theory with the SCM, we examined whether anthropomorphic design cues influence prosocial behavior (i.e., donations) by shaping perceptions of warmth and competence and reducing algorithm aversion. The results show that the anthropomorphic design of CAs increased both perceived warmth and perceived competence. However, only competence translated into downstream effects by reducing algorithm aversion and increasing donation amounts. In contrast, warmth did not significantly affect algorithm aversion or prosocial behavior.

Additional exploratory mediation analyses indicated indirect associations through perceived competence, but not through perceived warmth.

Effects of Anthropomorphizing on Warmth and Competence

Our findings suggest that anthropomorphic design cues shape how users socially evaluate CAs. Consistent with our theoretical integration, anthropomorphic cues appear to make the CA more socially legible by encouraging users to attribute human-like qualities to it. These attributions are then reflected in the core dimensions of the SCM, namely perceptions of warmth and competence. Thus, the design of anthropomorphic cues does not merely create a general impression of humanness but translates into specific judgments about the agent's intentions and capabilities.

The increase in warmth is theoretically intuitive, as warmth reflects benevolence, friendliness, and empathy, attributes closely linked to human nature (Abele & Wojciszke, 2007; Haslam, 2006). The increase in competence is more theoretically informative. While machines may already be associated with rationality and efficiency, competence in prosocial decision-making requires more than analytical processing. This corresponds with findings from Cheng (2022) and Schmid et al., (2022). For our research context, charity situations involve understanding values, integrating facts and feelings, and offering guidance that appears meaningful and morally appropriate. In this context, anthropomorphistic designs may allow the CA to appropriate qualities typically reserved for human actors. At the same time, the findings suggest that natural language alone may not be sufficient to meaningfully shift social evaluations. Rather, salient human-like cues appear necessary for users to perceive the CA as both warmer and more competent.

Effects of Warmth and Competence on Aversion and Prosocial Behavior

The central pattern in our results is that competence was the only dimension significantly associated with both lower algorithm aversion and higher donations, even though anthropomorphic design raised both dimensions. Even so paths did not differ significantly from each other, our results shade a different light on the primacy of warmth contradicting findings from other research (e.g., Kim et al., 2019; Gilad et al., 2021). Although an anthropomorphic design increased both dimensions, competence was the only predictor of lower algorithm aversion and higher prosocial behavior (i.e., donation amounts). This finding nuances the SCM in AI-mediated prosocial decision-making. While warmth is often considered primary in social and moral judgments in human-to-human interactions, users in our setting appeared to rely on whether the CA seemed capable of supporting an effective donation decision. From a theoretical perspective, this suggests that prosocial behavior in interactions with AI may depend on perceived ability to translate moral intentions into meaningful outcomes (Wang et al., 2016). In charity selection, users may already possess sufficient moral concern themselves. What they require from the CA is competent guidance, such as structuring information, comparing charities, and increasing confidence that the selected charity is suitable. The absence of a significant effect of algorithm aversion on prosocial behavior (i.e., donation amount) further indicates that negative evaluations of the CA do not automatically transfer to the prosocial object. Donation decisions may also reflect personal financial constraints, pre-existing charity preferences, or perceptions of the available charities. An alternative account of the donation results is social pressure: a more human-like counterpart may make declining to donate psychologically costlier (DellaVigna et al., 2012), a pathway our design cannot separate from the perceptual ones. Overall, competence appears to be the main mechanism linking anthropomorphic AI design to prosocial behavior. However, the high levels of perceived competence across all conditions suggest a potential ceiling effect. As LLM capabilities continue to advance, the scope for further increasing perceived competence through anthropomorphic design may become increasingly limited.

The Role of Attitude Towards AI as a Boundary Condition

We conducted additional robustness checks with attitude towards AI because this construct showed strong explanatory power and indicated potential heterogeneity in our model. When eliminating attitude towards AI from our model, the effect of warmth on algorithm aversion became significant. Rendering our results fragile. A deeper analysis of the boundary condition of attribute towards AI on our model's pathways, and particularly the role of warmth and competence will be an interesting avenue for future research. Thus, attitude toward AI may act as a boundary condition: anthropomorphism could shape social evaluations of the CA, but prior attitudes may influence whether these evaluations become behaviorally relevant.

Comparison of the Conversational Agents

The comparison across conditions provides a nuanced view of the value of CAs in prosocial decision-making. The computer-like CA was not perceived as warmer, more competent, or more anthropomorphized than the No-CA condition. This suggests that natural-language interaction alone may not be sufficient to generate the social-evaluative benefits often associated with CAs. Practically, this matters because a text-based interface is easier and cheaper to implement than an interactive CA. Unless the CA is sufficiently anthropomorphized, it may offer limited added value. At the same time, the human-like CA performed more favorably on several perceptual dimensions. It increased warmth and competence and was associated with higher enjoyment compared with the No-CA condition. Although anthropomorphism did not directly increase donation amounts, the mediation results suggest an indirect pathway through perceived competence. Thus, the human-like CA appears to provide stronger social-evaluative benefits, but these benefits are specific rather than reflecting a general improvement across all outcomes. Interestingly, both CA variants reduced algorithm aversion relative to the No-CA condition, while not differing from each other. For the human-like CA, this reduction can be explained through competence. For the computer-like CA, the mechanism remains unclear.

Limitation

Three overarching limitations bound our conclusions. First, the study design and measurement constrain causal specificity. The anthropomorphism manipulation combines verbal, identity, and nonverbal cues, the No-CA condition was fielded non-concurrently, the mediators were measured rather than manipulated, and the primary aversion measure relied on a single comparative item. Future studies could isolate individual cues through factorial designs, collect all conditions concurrently, manipulate the proposed mediators, and examine behavioral reliance.

Second, the study's context limits generalizability. We examine lottery-incentivized giving to animal charities among US online participants. Replications should test other populations, incentive structures, prosocial domains, and advice contexts, particularly to identify what mediates donation effects when perceptual pathways are inactive.

Third, the findings are bounded by the technological context at the time of data collection. CA capabilities and user expectations have evolved considerably (Susarla et al., 2023). Replications with current systems should examine whether increasing anthropomorphization will still achieve an increase in warmth or competence.

Contribution

Our findings contribute to research on the design of anthropomorphic CAs, algorithm aversion, and prosocial decision-making by showing that anthropomorphised CAs shape social evaluations, but that these evaluations are not equally consequential. While anthropomorphic cues increased both warmth and competence, only competence consistently reduced algorithm aversion and increased donation amounts. This is theoretically noteworthy because prosocial contexts, particularly donation contexts, are emotionally charged and would suggest a strong role of warmth (Zhou et al., 2019). Yet our findings indicate that users may already bring sufficient moral concern to the task and instead rely on the CA for effective, meaningful, and context-sensitive guidance.

Theoretically, this refines the SCM by showing that the primacy of warmth may not fully transfer to AI-mediated prosocial decision-making. By integrating SCM with self-humanization theory, we show that the anthropomorphism of CAs can help to incorporate appropriate human-like qualities, with competence as the dimension that becomes behaviorally actionable. This suggests that users may evaluate anthropomorphized technologies by whether they can translate prosocial intentions into suitable outcomes – even in prosocial settings. Practically, our results imply that designers of CAs for human-facing and morally relevant tasks should ensure that anthropomorphic design is paired with clear signals of competence, while also considering users' prior attitudes toward AI as a potential boundary condition. Furthermore, both CAs were more enjoyable than a text-only interface, which can foster favorable attitudes toward the provider (Araujo, 2018).

References

Aaker, J., Vohs, K. D., & Mogilner, C. (2010). Nonprofits Are Seen as Warm and For-Profits as Competent: Firm Stereotypes Matter. *Journal of Consumer Research*, *37*(2), 224–237. https://doi.org/10.1086/651566

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Ahn, J., Kim, J., & Sung, Y. (2022). The effect of gender stereotypes on artificial intelligence recommendations. *Journal of Business Research*, *141*, 50–59. https://doi.org/10.1016/j.jbusres.2021.12.007

Alashoor, T., Keil, M., Jiang, Z. (Jack), & Saffarizadeh, K. (2025). Privacy concerns and data donations: Do societal benefits matter? *MIS Quarterly*, *49*(2), 429–464. https://doi.org/10.25300/MISQ/2024/16853

Aragonés, J. I., Poggio, L., Sevillano, V., Pérez-López, R., & Sánchez-Bernardos, M.-L. (2015). Measuring warmth and competence at inter-group, interpersonal and individual levels. *Revista de Psicología Social*, *30*(3), 407–438. https://doi.org/10.1080/02134748.2015.1065084

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Baird, A., & Maruping, L. M. (2021). The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. *MIS Quarterly*, *45*(1), 315–341. https://doi.org/10.25300/MISQ/2021/15882

Banks, J. (2019). A perceived moral agency scale: Development and validation of a metric for humans and social machines. *Computers in Human Behavior*, *90*, 363–371. https://doi.org/10.1016/j.chb.2018.08.028

Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, *77*(3), 525–537. https://doi.org/10.1037/0022-3514.77.3.525

Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological Science*, *29*(5), 834–844. https://doi.org/10.1177/0956797617747648

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Blicher, A., Gleasure, R., Constantiou, I., & Clement, J. (2026). How do attention and emotion drive donations in charitable crowdfunding? An empirical study of users' cognitive and physiological responses to image content. *European Journal of Information Systems*, 1–22. https://doi.org/10.1080/0960085X.2026.2631718

Brendel, A. B., Hildebrandt, F., Dennis, A. R., & Riquel, J. (2023). The paradoxical role of humanness in aggression toward conversational agents. *Journal of Management Information Systems*, *40*(3), 883–913. https://doi.org/10.1080/07421222.2023.2229127

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Cervellon, M.-C., Poujol, J. F., & Tanner, J. F. (2019). Judging by the wristwatch: Salespersons' responses to status signals and stereotypes of luxury clients. *Journal of Retailing and Consumer Services*, *51*, 191–201. https://doi.org/10.1016/j.jretconser.2019.04.013

Cheng, L. (2022). The effects of smartphone assistants' anthropomorphism on consumers' psychological ownership and perceived competence of smartphone assistants. *Journal of Consumer Behaviour*, *21*(2), 427–442. https://doi.org/10.1002/cb.2021

Christoforakos, L., Gallucci, A., Surmava-Große, T., Ullrich, D., & Diefenbach, S. (2021). Can robots earn our trust the same way humans do? A systematic exploration of competence, warmth, and anthropomorphism as determinants of trust development in HRI. *Frontiers in Robotics and AI*, *8*, 640444. https://doi.org/10.3389/frobt.2021.640444

Chu, K., Lee, D.-H., & Kim, J. Y. (2016). The effect of non-stereotypical gender role advertising on consumer evaluation. *International Journal of Advertising*, *35*(1), 106–134. https://doi.org/10.1080/02650487.2015.1110942

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149). Elsevier. https://doi.org/10.1016/S0065-2601(07)00002-0

De Wit, A., & Bekkers, R. (2016). Exploring Gender Differences in Charitable Giving: The Dutch Case. *Nonprofit and Voluntary Sector Quarterly*, *45*(4), 741–761. https://doi.org/10.1177/0899764015601242

DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, *127*(1), 1–56. https://doi.org/10.1093/qje/qjr050

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*(1), 91–119. https://doi.org/10.1037/0033-2909.101.1.91

El Hedhli, K., Zourrig, H., Al Khateeb, A., & Alnawas, I. (2023). Stereotyping human-like virtual influencers in retailing: Does warmth prevail over competence? *Journal of Retailing and Consumer Services*, *75*, 103459. https://doi.org/10.1016/j.jretconser.2023.103459

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

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, *132*, 138–161. https://doi.org/10.1016/j.ijhcs.2019.07.009

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

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 HCI*, *35*(6), 456–467. https://doi.org/10.1080/10447318.2018.1456150

Gefen, D., & Straub, D. (2003). Managing user trust in B2C e-services. *E-Service Journal*, *2*(2), 7–24. https://doi.org/10.2979/esj.2003.2.2.7

Gilad, Z., Amir, O., & Levontin, L. (2021). The effects of warmth and competence perceptions on users' choice of an AI system. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3411764.3446863

Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Does PLS Have Advantages for Small Sample Size or Non-Normal Data? *MIS Quarterly*, *36*(3), 981–1001. https://doi.org/10.2307/41703490

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making*, *26*(3), 213–224. https://doi.org/10.1002/bdm.1753

Grassini, S. (2023). Development and validation of the AI attitude scale (AIAS-4): A brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology*, *14*, 1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of Mind Perception. *Science*, *315*(5812), 619–619. https://doi.org/10.1126/science.1134475

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)* (1st edn). SAGE Publications Inc.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Heßler, P. O., Pfeiffer, J., & Hafenbrädl, S. (2022). When self-humanization leads to algorithm aversion. *Business & Information Systems Engineering*, *64*(3), 275–292. https://doi.org/10.1007/s12599-022-00754-y

Huang, Z., Che, C., Zheng, H., & Li, C. (2024). Research on Generative Artificial Intelligence for Virtual Financial Robo-Advisor. *Academic Journal of Science and Technology*, *10*(1), 74–80. https://doi.org/10.54097/30r2kk80

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. *In Proceedings of the 28th European Conference on Information Systems (ECIS)*. https://doi.org/https://aisel.aisnet.org/ecis2020_rp/168

Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T., & Bernstein, M. S. (2020). Conceptual metaphors impact perceptions of human-AI collaboration. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1–26. https://doi.org/10.1145/3415234

Kim, S. Y., Schmitt, B. H., & Thalmann, N. M. (2019). Eliza in the uncanny valley: Anthropomorphizing consumer robots increases their perceived warmth but decreases liking. *Marketing Letters*, *30*(1), 1–12. https://doi.org/10.1007/s11002-019-09485-9

Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can Machines Think? Interaction and Perspective Taking with Robots Investigated via fMRI. *PLoS ONE*, *3*(7), e2597. https://doi.org/10.1371/journal.pone.0002597

Li, Q., Luximon, Y., & Zhang, J. (2023). The Influence of Anthropomorphic Cues on Patients' Perceived Anthropomorphism, Social Presence, Trust Building, and Acceptance of Health Care Conversational Agents: Within-Subject Web-Based Experiment. *Journal of Medical Internet Research*, *25*, e44479. https://doi.org/10.2196/44479

Li, Y., Gan, Z., & Zheng, B. (2025). How do artificial intelligence chatbots affect customer purchase? Uncovering the dual pathways of anthropomorphism on service evaluation. *Information Systems Frontiers*, *27*(1), 283–300. https://doi.org/10.1007/s10796-023-10438-x

Liao, J., & Huang, J. (2024). Think like a robot: How interactions with humanoid service robots affect consumers' decision strategies. *Journal of Retailing and Consumer Services*, *76*, 103575. https://doi.org/10.1016/j.jretconser.2023.103575

Loewenstein, G., & Small, D. A. (2007). The scarecrow and the tin man: The vicissitudes of human sympathy and caring. *Review of General Psychology*, *11*(2), 112–126. https://doi.org/10.1037/1089-2680.11.2.112

Logg, J. M. (2017). Theory of Machine: When Do People Rely on Algorithms? *Harvard Business School Working Paper Series # 17-086*. https://dash.harvard.edu/handle/1/31677474

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, *25*(3), 695–710. https://doi.org/10.1016/j.chb.2008.12.026

Moon, J., & Kim, Y.-G. (2001). Extending the TAM for a World-Wide-Web context. *Information & Management*, *38*(4), 217–230. https://doi.org/10.1016/S0378-7206(00)00061-6

Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research*, *26*(4), 323–339. https://doi.org/10.1086/209566

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLoS ONE*, *9*(12), e115844. https://doi.org/10.1371/journal.pone.0115844

Prahl, A., & van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, *36*(6), 691–702. https://doi.org/10.1002/for.2464

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *36*(4), 717–731. https://doi.org/10.3758/BF03206553

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people.*

Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4*. SmartPLS. https://www.smartpls.com/

Rossignac-Milon, M., Bolger, N., Zee, K. S., Boothby, E. J., & Higgins, E. T. (2021). Merged minds: Generalized shared reality in dyadic relationships. *Journal of Personality and Social Psychology*, *120*(4), 882–911. https://doi.org/10.1037/pspi0000266

Schanke, S., Burtch, G., & Ray, G. (2021). Estimating the impact of "humanizing" customer service chatbots. *Information Systems Research*, *32*(3), 736–751. https://doi.org/10.1287/isre.2021.1015

Schmid, D., Staehelin, D., Bucher, A., Dolata, M., & Schwabe, G. (2022). Does social presence increase perceived competence?: Evaluating conversational agents in advice giving through a video-based survey. *Proceedings of the ACM on Human-Computer Interaction*, *6*(GROUP), 1–22. https://doi.org/10.1145/3492845

Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems*, *37*(3), 875–900. https://doi.org/10.1080/07421222.2020.1790204

Seeger, A.-M., Pfeiffer, J., & Heinzl, A. (2021). Texting with humanlike conversational agents: Designing for anthropomorphism. *Journal of The Association for Information Systems*, *22*(4), 931–967. https://doi.org/10.17705/1jais.00685

Seymour, M., Yuan, L. (Ivy), Riemer, K., & Dennis, A. R. (2025). Less artificial, more intelligent: Understanding affinity, trustworthiness, and preference for digital humans. *Information Systems Research*, *36*(2), 1096–1128. https://doi.org/10.1287/isre.2022.0203

Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems. *Information Systems Research*, *34*(2), 399–408. https://doi.org/10.1287/isre.2023.ed.v34.n2

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Tinwell, A., & Sloan, R. J. S. (2014). Children's perception of uncanny human-like virtual characters. *Computers in Human Behavior*, *36*, 286–296. https://doi.org/10.1016/j.chb.2014.03.073

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Wang, A., & Rhemtulla, M. (2021). Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920918253. https://doi.org/10.1177/2515245920918253

Wang, Z., Mao, H., Jessica Li, Y., & Liu, F. (2016). Smile big or not? Effects of smile intensity on perceptions of warmth and competence. *Journal of Consumer Research*, (43), 787–805. https://doi.org/10.1093/jcr/ucw062

Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology*, *38*(7), 1139–1147. https://doi.org/10.1002/ejsp.549

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the Dominance of Moral Categories in Impression Formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251–1263. https://doi.org/10.1177/01461672982412001

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research*, *37*(2), 197–206. https://doi.org/10.1086/651257

Zhou, X., Kim, S., & Wang, L. (2019). Money helps when money feels: Money anthropomorphism increases charitable giving. *Journal of Consumer Research*, *45*(5), 953–972. https://doi.org/10.1093/jcr/ucy012