European Journal of Marketing

55

The impact of providing non-human identity cues about sales agents on consumer responses: the role of social presence and speciesism activation

Roberta De Cicco
Department of Neuroscience, Imaging and Clinical Sciences, Gabriele d'Annunzio University of Chieti and Pescara, Pescara, Italy, and Department of Communication Sciences, Humanities and International Studies (DISCUI), University of Urbino "Carlo Bo", Urbino, Italy

Maher Georges Elmashhara
Research Centre in Management and Economics, Católica Porto Business School, Universidade Católica Portuguesa, Porto, Portugal, and Manchester Metropolitan University Business School, Manchester, UK

Susana C. Silva
Católica Porto Business School and Research Centre in Management and Economics, Universidade Católica Portuguesa, Porto, Portugal, and Faculty of Business and Law, University of Saint Joseph, Macao SAR, China, and

Maik Hammerschmidt
Department of Marketing and E-Business, Retail Logistics & Innovation Lab, University of Goettingen, Goettingen, Germany

Received 28 January 2022
Revised 16 November 2022
25 May 2023
27 October 2023
21 April 2024
20 November 2024
27 February 2025
Accepted 2 March 2025

Abstract

Purpose – This work investigates how different strategies for providing cues about the non-human identity of a sales agent influence consumers' perceptions and purchase-related outcomes, and how a social interaction style shapes these responses. Additionally, the authors explore the role of consumers' speciesism against non-human entities in eliciting unfavourable responses to the disclosure of the agent's artificial nature.

Open Access funding enabled and organised by Projekt DEAL.

Funding: The third author is funded by Fundação para a Ciência e Tecnologia (through project UIDB/00731/2020).

Conflict of interest/competing interests: The authors declare that they have no conflict of interest, no competing financial interests exist.

Data and code availability: Available at reasonable request.

European Journal of Marketing
Vol. 59 No. 13, 2025
pp. 55-84
Emerald Publishing Limited
0309-0566
DOI 10.1108/EJM-01-2022-0066

Design/methodology/approach – Three experimental studies were conducted using real chatbot interactions. Study 1 investigates how non-human identity cues impact consumer trust and, subsequently, attitude towards the firm and intention to purchase the product offered. Study 2 tests these effects across different levels of social presence. Study 3 examines consumer responses to different non-human identity disclosure strategies, considering speciesism's moderating role.

Findings – Study 1 proves that disclosing (vs not disclosing) the artificial nature of a sales agent leads to a decline in trust towards the firm, which in turn negatively influences both attitude towards the firm and purchase intention. This finding reveals discrimination against disclosed (vs non-disclosed) artificial sales agents despite identical, flawless performance. However, Study 2 proves that the negative effects vanish when perceived social presence is high. Study 3 underlines that high speciesism leads to a trust decline if non-human identity cues are presented during the interaction but not if presented earlier in the journey before the interaction.

Research limitations/implications – The study highlights the negative effects of disclosure on important, firm-related outcomes. These insights advance current literature by showing that disclosing cues about the non-human nature of a sales agent can undermine psychological and behavioural responses–even when the disclosed agent performs just as effectively as its undisclosed counterpart. This result is noteworthy, as most prior research has linked aversive reactions to artificial agents with situations in which algorithms underperform, whereas this study examines agents that function flawlessly. Furthermore, the study reveals that these adverse effects are driven by speciesism–prejudices against non-human entities–offering a novel explanation for consumers' negative responses.

Practical implications – The findings stress that transparency about the artificial nature of sales agents is penalised by customers and comes at a high cost for business-relevant outcomes. However, by transforming an artificial agent into a social actor through subtle design modifications, firms can overcome the unfavourable prejudice against artificial agents. By creating a social appearance, firms can harness the potential of automated sales services–even when disclosure of the agent's artificial identity is required. As firms may soon be obliged to disclose the artificial identity of their sales agents, the critical question shifts from whether to disclose to how to disclose in order to mitigate negative consequences. Finally, we offer guidance on targeting the right consumers with artificial agents–specifically, those with lower levels of speciesism-related prejudices.

Originality/value – This work addresses pressing issues for managers concerned with the implementation of artificial sales agents. Results extend knowledge on speciesism towards digital agents, inform which consumers are particularly prone to respond negatively to such agents, and present levers for designing chat-based social interactions that prevent non-human-related prejudices that could undermine the effectiveness of conversational technologies.

Keywords Artificial agents, Chatbots, Sales agents, Identity disclosure, Social presence, Speciesism, Trust, Attitude, Conversational commerce

Paper type Research paper

1. Introduction

Artificial agents, such as chatbots, are programs designed to emulate aspects of human intelligence by simulating human conversations (Chandra *et al.*, 2022; Murtarelli *et al.*, 2021), making them increasingly popular across social media platforms, online interfaces and mobile apps. These agents have been adopted across a broad range of consumer domains (Schuetzler *et al.*, 2018). They can execute simple tasks, such as offering information or sending airline tickets, or more complex ones, such as giving shopping recommendations and selling products (Garvey *et al.*, 2022; Soares *et al.*, 2022). It is no surprise that artificial sales agents are expected to become the dominant interface for purchase transactions in many settings (Davenport *et al.*, 2020; Thomaz *et al.*, 2020). As businesses increasingly integrate these agents into customer interactions, it is important to explore when and how to disclose the artificial nature of such agents, how this disclosure influences consumer responses, and which consumer characteristics are relevant in shaping these responses.

For consumers, artificial sales agents offer a range of benefits and opportunities, including the ability to facilitate round-the-clock transactions and responses to inquiries (Bozkurt *et al.*, 2020). Beyond the higher speed of service provision, they can deliver consumers the right offers at the right time and provide a wider variety and quantity of purchase-related

information compared with traditional agents, allowing for more efficient transactions (Huang and Rust, 2018). While numerous studies suggest that consumers exhibit improved acceptance of automated service provision in some domains when using specific cues, such as affective human-like qualities (Castelo *et al.*, 2019) or features like increased personalisation (Longoni *et al.*, 2019), other research highlights potential consumer aversion when an agent reveals its artificial nature during interactions (Luo *et al.*, 2019). Additionally, users might distrust computer programs to reveal personal needs or make purchase decisions because they perceive artificial agents as less empathetic and knowledgeable (Mozafari *et al.*, 2020) and thoroughly prefer humans unless artificial agents perform notably superior, especially in the customer service domain (Castelo *et al.*, 2023).

Given these opposing arguments, extant research has not reached a consensus o*n whether* companies that employ digital agents should provide explicit cues about their artificial identity to consumers (Mozafari *et al.*, 2022; Murtarelli *et al.*, 2021; Van der Goot *et al.*, 2024). This issue becomes particularly crucial as artificial tools (like chatbots) can emulate human-like interactions and could be easily mistaken for humans if their artificial nature is not disclosed. For example, Zalando recently launched a fashion chatbot that allows customers to ask questions in their own words, with responses provided intuitively and naturally, similar to those of human employees (Retail Week, 2023). Research is also silent o*n how* the non-human identity should be disclosed in a way that minimises potential harmful effects on trust and attitude towards the firm and subsequent key outcomes like purchase. This is a severe research gap, as firms might soon be legally forced to disclose the artificial identity and must find the right way to disclose it through proper interaction design, especially in light of the game-changing role of Generative AI tools that are recasting the way not only business but many daily life routines are performed (Peres *et al.*, 2023). Another open question is for *whom* a disclosure has particularly devastating effects. This gap is closely connected to the question of which consumers display a pronounced prejudice against non-human entities like artificial sales agents (Liu-Thompkins *et al.*, 2022; Schmitt, 2020).

Based on these gaps regarding whether disclosing the artificial nature of an agent has negative effects on business outcomes, how artificial agents should be designed and to whom a disclosure should be communicated, the following research questions arise: (1) Does disclosing (vs not disclosing) the artificial identity of an agent to consumers generate a spillover effect undermining trust towards the firm and, consequently, attitude towards the firm and purchase intention? (2) Can eventual negative responses be countered through appropriate social-oriented interaction design? (3) Are eventual negative responses contingent on personal prejudices towards non-human species?

To answer these questions, we conduct three studies using real-life, synchronous chatbot interactions and analyse the effects of different strategies of non-human identity disclosure of a sales agent (no disclosure, disclosure before the interaction and disclosure within the interaction) on downstream consequences for firms. Additionally, we explore specific conditions (low vs high social presence of the digital agent and low vs high speciesism of its users) under which these effects occur.

The findings of Study 1 highlight the negative effects of disclosure on trust towards the firm and, indirectly, on attitude towards the firm and intention to purchase. These insights advance current literature by showing that disclosing cues about the artificial nature of an agent during an interaction harms psychological and purchase-related responses, although the disclosed non-human agent performs equally well as the undisclosed counterpart. This result is especially noteworthy as most research has documented aversive reactions to artificial agents mainly for settings where algorithms fail (e.g. Dietvorst *et al.*, 2015), while

we consider agents that perform flawlessly. Our observation reveals that a trust decrease explains these harmful effects, although no human employees are involved. This is a novel insight, as extant studies have primarily focused on comparing trust towards artificial vs human agents (Garvey *et al.*, 2022; Longoni *et al.*, 2019). Additionally, ours is one of the first studies to explore the negative effects of artificial identity disclosure on the firm that employs the digital agent. In doing so, we uncover a negative spillover effect of artificial agents on evaluations of the company itself. Interestingly, the significant trust-eroding effect of providing non-human identity cues prevailed even if participants knew beforehand that they were interacting with an artificial agent. Reminding people about the known artificial identity of the agent elicits a trust reduction towards the firm compared to identical interactions without such cues. Our findings stress that transparency about the artificial nature of the conversational agent is penalised by consumers and comes at a high cost for business-relevant outcomes.

Based on the need to identify ways to reduce this negative effect found in Study 1, in Study 2 we show that when the social presence of the conversational agent is low–that is, when linguistic and pragmatic cues are poor–undesirable responses to the disclosure increase. In contrast, when making a social actor out of an artificial agent through slight design modifications that integrate social elements, the variation in trust disappears. These results provide important implications for settings where the question is no longer only *whether* to disclose the non-human identity but *how* to disclose it to be transparent and avoid harmful effects. We show that designing sociable agents is an effective managerial intervention for avoiding a trust discount, even if their artificial identity needs to be disclosed.

In Study 3, we reveal that the decrease in trust after the identity disclosure mainly occurs for consumers with high levels of speciesism–a belief in the human species' superiority over non-human species (Caviola *et al.*, 2019; Schmitt, 2020). In contrast, consumers with low speciesism do not exhibit a trust decline and hence no undesirable responses to disclosed agents. These results indicate that the activation of a speciesism bias against non-human entities is an essential explanation for the trust erosion after the artificial identity disclosure. Interestingly, the results further reveal that the instigation of speciesism occurs only if non-human identity cues are provided *within* the interaction, while for disclosure *before* the interaction, no speciesism activation and hence no negative effects can be observed. These results highlight that speciesism can be prevented through the right timing of non-human identity disclosure in the purchase journey.

2. Research framework and hypotheses development

2.1 Research framework

To evaluate whether, how, and for whom openly disclosing the artificial nature of a sales agent (chatbots in our case) yields unfavourable firm consequences, we consider the effects of non-human identity disclosure (vs no disclosure) on attitude towards the firm and intention to purchase the offered product through trust towards the firm. We also delve deeper into different strategies of non-human identity disclosure by examining the effects of the disclosure *before* the interaction vs disclosure *within* the interaction. We also examine the consequences of identity disclosure for different levels of social presence and consumer speciesism against non-human entities as moderators.

Trust represents the willingness to rely on and be vulnerable to the other party in whom one has confidence and expectations of fairness (Castaldo *et al.*, 2010). In the context of new technologies, trust is a crucial metric as consumers need to be able to trust that firms will help

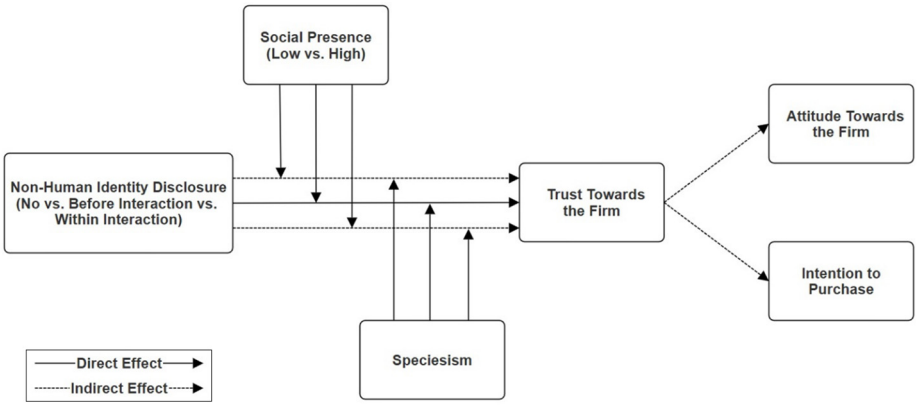
them overcome perceptions of risk in the use of new technologies (Li *et al.*, 2008) and make good use of their data (Kaplan and Haenlein, 2020).

Thus, we use trust towards the firm as the mediator–rather than trust towards the agent as in previous research (e.g. Przegalinska *et al.*, 2019)–to identify possible spillover effects from the digital agent to the firm. Since the firm has decided to employ a digital agent for customer interaction, its trust evaluations might be influenced by this decision. We selected attitude towards the firm and intention to purchase a product from the firm as outcome metrics, as these are key indicators of actual usage and purchase behaviour. This is grounded in the idea that future actions are shaped by an individual's cognitive evaluations and their intentions to perform specific tasks (Morwitz *et al.*, 2007; Sheeran, 2002). Attitude, which represents a tendency to respond positively or negatively towards a certain object or person (Fishbein and Ajzen, 1977), plays a crucial role in shaping observable purchase activities (Schellekens *et al.*, 2010). Similarly, recent studies in the digital agent context empirically confirm that the intention to purchase can be considered a reliable predictor for actual purchases (Elmashhara *et al.*, 2024; Holthöwer and Van Doorn, 2023). Since attitude and purchase intention refer to different objects (firm and product, respectively), both variables are included as distinct outcomes, in line with extant studies (e.g. Evans *et al.*, 2017; Van Reijmersdal *et al.*, 2022).

In distinguishing our independent variable (artificial identity disclosure) and moderator variables (social presence, speciesism), we emphasise that disclosure refers to a communication strategy of the firm (namely whether to communicate the artificial nature of the agent irrespective of its design) while our moderator social presence reflects design features of the sales agent that can be manipulated by the firm (like social cues) and speciesism captures a consumer characteristic relevant for segmentation and targeting. Figure 1 summarises our research framework.

2.2 Discriminating between artificial agents: the effects of non-human identity disclosure on trust towards the firm

At first glance, being honest about the non-human identity of the conversational partner should enhance transparency and ensure that consumers do not feel exploited by firms



Source(s): Authors' own work

Figure 1. Research framework

through a lack of identity information (Følstad *et al.*, 2018). Artificial interfaces have advanced to the point where they can mimic human conversations so convincingly that consumers confuse digital agents with human assistants. This is particularly evident with disruptive Generative AI tools like ChatGPT (Peres *et al.*, 2023). As practices for increasing transparency inevitably remind users about the machine identity of a conversational partner, recent literature suggests that this comes with highly unfavourable consequences for trust. Studies on algorithm aversion show that people distrust algorithms if they see them making the same mistake as human employees (e.g. providing a suboptimal recommendation), even if they know that algorithms will outperform humans in an increasingly comprehensive list of service tasks from diagnosing diseases to giving legal advice (Castelo *et al.*, 2019; Luo *et al.*, 2019).

Research has yet to deeply examine whether the aversion against artificial agents occurs even if no human employees are involved (i.e. only non-human agents with different identities are considered) and if the agents perform error-free. We argue that the effect of algorithms on trust towards the firm not only prevails when compared to human interactions and in failure situations but even occurs for disclosed algorithms compared to non-disclosed algorithms with the same capabilities and providing identical performance. Thus, besides offering an important extension on the crucial "disclosure dilemma", which advocates that disclosing the machine identity of an agent to consumers might impair the business value of digital agents because of their pushback (Mozafari *et al.*, 2020), our study emphasises the risk of a spillover effect. This spillover effect arises when suspicion towards the artificial agent distorts the trust evaluations of the entire company (Sangle-Ferriere and Voyer, 2019). Hence, we extend knowledge on a specific manifestation of trust by theorising about differences in trust towards the firm based on different disclosure cues for artificial sales agents when no human alternative exists.

Specifically, we suggest that delivering artificial intelligence cues acts as an information priming on users (Iacobucci *et al.*, 2021; Roskos-Ewoldsen and Fazio, 1992). By including non-human cues in an interaction, people may associate the experience with common concerns about recent conversational technologies, such as voice assistants or smart speakers (e.g. concerns about privacy invasion by companies and uncertainty around how such devices bias firm decisions), which then influence their evaluation of the firm (Uysal *et al.*, 2022). Consequently, making the unfavourable cognitions salient for identity-disclosed artificial agents erodes trust in the firm compared to an undisclosed agent, even if both perform flawlessly and are perfect service agents. The actual performance of the identified artificial agent might not matter in forming trust evaluations if the non-human identity cues assigned to the agent have already elicited unfavourable judgments. This would not be the case if people had no information on the nature of the conversational agent. Based on that, we hypothesise that the disclosure of the agent's (e.g. chatbot's) artificial identity negatively affects trust towards the firm deploying the agent:

H1. Non-human identity disclosure (vs no disclosure) reduces trust towards the firm.

2.3 No trust, no transaction: the mediating role of trust towards the firm

Trust has been extensively studied as a critical element in the relationships between individuals and organisations (Hong and Cha, 2013; Suh and Han, 2003). In e-commerce settings, trust in a firm plays a pivotal role in shaping both consumer attitude (Kim and Peterson, 2017) and purchase intention (Qureshi *et al.*, 2009). Regarding attitude, trust positively influences consumers' perceptions of the organisation because it fosters confidence in the firm's reliability and ability to deliver on its promises (Kim and Peterson, 2017).

This confidence builds a sense of assurance and alignment with the organisation's values, leading to a more favourable attitude towards the firm that employs a conversational agent (De Cicco *et al.*, 2020). Regarding purchase intention, trust enhances the consumer's willingness to engage in transactions by reducing feelings of uncertainty and hesitation (Pavlou, 2003). Consumers are more likely to commit to purchasing when they believe the firm will uphold its commitments and deliver quality products or services (Gu *et al.*, 2021). Moreover, trust instils a sense of reliability in the firm's offerings, reinforcing confidence in the purchasing process (Aydin and Özer, 2005; Wong and Haque, 2022).

Under the notion that trust towards the firm positively affects attitudes and purchase intentions, if non-human identity disclosure reduces trust, attitudinal and purchase-related outcomes will indirectly be affected negatively. In this case, we believe trust mediates the relation between the interaction with the digital agent and the firm-relevant outcomes. Based on this rationale, we hypothesise:

H2a. Non-human identity disclosure has a negative indirect effect on attitude towards the firm via trust towards the firm.

H2b. Non-human identity disclosure has a negative indirect effect on intention to purchase via trust towards the firm.

2.4 "AI feel you": the moderating role of social presence

As argued earlier, the anticipated decrease in trust might be driven by the belief that artificial agents lack personal feelings and empathy. Hence, designing human interactions with artificial agents to eliminate this belief should be a powerful lever for reducing negative feelings induced by the artificial identity cue, thus building trust with the firm deploying the agent. This aligns with research highlighting the vital role of anthropomorphism in fostering trust in situations involving digital agents (Munnukka *et al.*, 2022). Social cues (e.g. the use of humour, a clear tone of voice, simple language and visual elements) are relevant in forming perceptions of interlocutors and evaluating the risk of stressful interactions (Murtarelli *et al.*, 2021). Drawing on social presence theory (Short *et al.*, 1976), we posit that cues that enhance the feeling of social presence, like anthropomorphising elements (Blut *et al.*, 2021; Castelo *et al.*, 2019) infused in human-agent interactions, are particularly potent to mitigate the unfavourable effects of the artificial identity disclosure on trust.

Social presence reflects the extent to which a medium is perceived to convey the feeling of human warmth and sociability and is found to facilitate media acceptance partly because of the human's desire for psychological connections with other humans (Munnukka *et al.*, 2022). Since humans articulate social expectations also to machines, in machine interactions, paralinguistic (e.g. GIF, emoticons) and pragmatic cues (e.g. personalised greetings such as calling the user by name or personalised answers such as "*great choice, we have the same taste*") play a major role in defining social traits (De Cicco *et al.*, 2021; Shawar and Atwell, 2005). In fact, infusing a digital agent, like a chatbot, with socially rich content should be particularly effective as it strongly fits the informal and easy-going nature of chats (De Cicco *et al.*, 2020).

Given these arguments, an increased feeling of intimacy through social presence intensifies the perceived warmth of the message source and, in turn, makes message evaluation more favourable (Van Doorn *et al.*, 2017). This should also happen with the message that discloses the non-human identity conveyed by the agent: social presence should positively shape the effect of identity disclosure on trust towards the firm. In fact, attributing social-emotional abilities and a human-like mind to a non-human agent makes consumers

consider the firm that employs such an agent more trustworthy due to consumers' tendency to overuse human social categories (Chandra *et al.*, 2022; Pizzi *et al.*, 2021). If an agent conveys a vivid and warm interaction, users should be able to cope with the fact that the conversational partner is non-human. High social presence makes users feel emotionally connected to the firm when using an agent (Steinhoff *et al.*, 2019), while the contrary should happen in the case of low social presence. Hence, we posit that social presence interacts with non-human identity disclosure in such a way that:

H3. Non-human identity disclosure has a less negative effect on trust towards the firm when social presence is high (vs low).

Finally, combining the argumentation that the higher social presence in a conversational system effectively mitigates the trust-eroding effects of non-human identity disclosure, and considering the mediating role of trust in the relationship between the identity disclosure and outcomes, we hypothesise that social presence moderates the mediated effect of the identity disclosure on outcomes via trust. Hence:

H4a. Social presence positively moderates the indirect effect of non-human identity disclosure on attitude towards the firm via trust towards the firm. That is, the indirect effect of non-human identity disclosure on attitude is less negative when social presence is high.

H4b. Social presence positively moderates the indirect effect of non-human identity disclosure on intention to purchase via trust towards the firm. That is, the indirect effect of non-human identity disclosure on intention to purchase is less negative when social presence is high.

2.5 "Artificial agents count for less than human?": the moderating role of speciesism

When elaborating on which people are particularly prone to negatively responding to digital agents in terms of trust decline, it is worth considering specific personality characteristics that could be relevant (Kopalle *et al.*, 2021). Specifically, we propose that a deeply ingrained prejudice regarding the worth of humans compared to non-human entities is particularly relevant in our study context, where tasks traditionally performed by humans are now being handled by machines. Research in philosophy and biology refers to such prejudice as speciesism–the belief that humans are intrinsically more valuable than members of other species (Ryder, 2006; Singer and Mason, 2007), such as animals (Caviola *et al.*, 2019). As speciesism is a psychological characteristic of individuals reflecting how they assign worth to individuals based on species membership, it can also be applied to interactions with mechanised actors as a different type of non-human species beyond animals (Schmitt, 2020).

In drawing upon this generalised interpretation of speciesism, we define speciesism as the extent to which individuals believe that non-human actors (such as artificial agents) are inferior to humans. In this sense, speciesism reflects that some people place the human species preferentially over non-human actors irrespective of their actual performance due to a prejudice that non-human actors lack cognitive abilities (Schmitt, 2020).

It has to be noted that speciesism goes beyond weirdness (Skjuve *et al.*, 2021), eeriness (Mende *et al.*, 2019) and algorithm aversion (Dietvorst *et al.*, 2015; Fuegener *et al.*, 2022). Although all constructs reflect uncomfortable conceptions towards technologies, the source of discomfort fundamentally differs: while weirdness or eeriness reflects a mismatch between the non-human and human qualities (Chen *et al.*, 2021; Mende *et al.*, 2019), and algorithm aversion results from the threat to human identity through technologies

encroaching on services that used to be provided by humans (Dietvorst *et al.*, 2015; Kopalle *et al.*, 2021), in speciesism this discomfort is produced by the fact that humans feel superior (Liu-Thompkins *et al.*, 2022; Schmitt, 2020). So basically, individuals may exhibit a negative response towards artificial agents not because they feel threatened or alienated, but because they believe they are superior to them based on a categorical comparison of humans and machines.

Based on the reasoning that speciesism is a latent personality characteristic of some individuals leading to conceptions of groups that diverge from reality, we choose this consumer characteristic to evaluate digital interactions that contain information about an artificial "out-group" (artificial agents, in our case). When people with high speciesism levels are confronted with artificial stimuli, they might automatically activate their pre-existing prejudice against machines, which might bias their evaluation of the interaction even in the absence of actual performance differences between groups. This will not be the case if people do not hold pronounced levels of speciesism. The activation of speciesism, therefore, also serves as a mechanism explaining the weakening of trust following disclosure.

However, we suggest that the timing of disclosing the non-human cues strongly determines whether speciesism prejudices are activated or not. First, it has to be noted that the level of speciesism should be particularly relevant if disclosure occurs, but less so in a no-disclosure setting, as in this setting, no artificial identity cues are communicated that could activate a speciesism bias. In the case of disclosure, we argue that presenting non-human identity cues of the agent *within* the interaction is much more severe than informing people about the non-human identity *before* the start of the interaction. If individuals with high latent speciesism are recalled they are interacting with a member of the devalued group (i.e. digital agents) through intra-conversational cues, this interruption of the conversation will prompt individuals to use their speciesism conceptions to evaluate the following flow of information (Bargh *et al.*, 1996). Thus, presenting non-human cues during a text-based interaction makes consumers' prejudice against machines highly accessible and hinders the feeling of being in a human-like interaction.

The opposite scenario occurs when the information about the artificial nature of the agent is decoupled from the conversation, such as when it is provided in the form of an initial disclaimer before starting the interaction, similar to how we often automatically accept cookies' releases. In that case, a speciesism bias is less likely to influence the judgment of the interaction, and hence a trust decline should be less pronounced. It is like racism: if an individual is not exposed to a specific interaction cue that activates racism, that prejudice should not become salient. In other words, we posit that there are different trust responses to artificial sales agents that introduce themselves *before* the interaction vs those that do so *within* the interaction due to different salience of speciesism related to these two strategies of non-human identity disclosure. Based on the argument mentioned above, we hypothesise:

H5. Within-interaction disclosure (vs before-interaction disclosure) has a more negative effect on trust towards the firm when speciesism is high compared to when speciesism is low.

So far, we have argued that high speciesism in users' minds becomes relevant when there is a need for an artificial identity disclosure, while for no disclosure, the speciesism level should be less relevant. In addition, we have argued that the timing of the disclosure (within- vs before interaction) matters for the moderating impact of speciesism regarding trust to unfold. Combining these arguments and the mediating role of trust in the relationship between non-human identity disclosure and outcomes, we hypothesise that speciesism moderates the

mediated effect of non-human identity disclosure on outcomes via trust. Therefore, the following hypotheses are proposed:

H6a. Speciesism negatively moderates the indirect effect of within-interaction (vs before-interaction) identity disclosure on attitude towards the firm via trust towards the firm. That is, the indirect effect of non-human identity disclosure on attitude is more negative when speciesism is high.

H6b. Speciesism negatively moderates the indirect effect of within-interaction (vs before-interaction) identity disclosure on intention to purchase via trust towards the firm. That is, the indirect effect of non-human identity disclosure on intention to purchase is more negative when speciesism is high.

2.6 Overview of studies

To test our hypotheses, three experiment-based studies have been conducted. In Study 1, we investigate the effect of disclosing the non-human identity of a conversational agent within the interaction on trust towards the firm (*H1*) and how this trust mediates the effect of this disclosure on attitude towards the firm and purchase intention (*H2a* and *H2b*). Study 2 examines the moderating effect of social presence in mitigating the negative effect of disclosure on trust (*H3*) and consequently on attitude and purchase intention (*H4a* and *H4b*). While Studies 1 and 2 consider the disclosure of artificial cues within the interaction and compare it with no disclosure, in Study 3, we differentiate between the timing of the non-human identity disclosure and, in addition to disclosure *within* the interaction, we consider a disclosure *before* the interaction. In doing so, Study 3 explores whether consumers with higher levels of speciesism trust the firm employing the artificial agent less if the non-human identity is disclosed within the interaction compared to earlier in the purchase journey (*H5*). In addition, we explore how speciesism moderates the mediating effect of trust for the relationships between the disclosure timing, and attitude and purchase intention (*H6a* and *H6b*).

3. Study 1

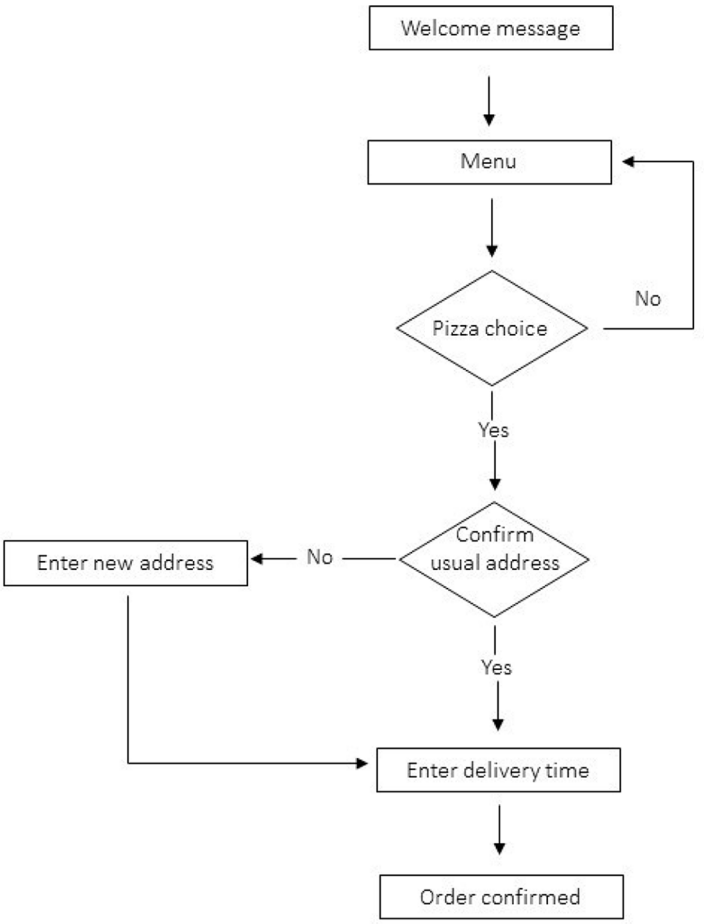
3.1 Design and experimental procedure

The goal of Study 1 is to examine how disclosing the non-human identity of a conversational agent impacts consumer trust (*H1*) and, in turn, attitude towards the firm and purchase intention (*H2a, H2b*). The study adopts a single-factor experimental design (no disclosure vs non-human identity disclosure). We recruited participants by sharing the link to the experiment on Facebook and LinkedIn and encouraged respondents to share the survey link among their acquaintances. The data was collected using Qualtrics, which allows randomly sending participants to either the control or experimental groups and places a cookie on participants' browsers when they submit a response to prevent multiple submissions.

A final sample of 160 European participants took part in the study: 77 participants in the no-disclosure condition and 83 participants in the disclosure condition. Participants ranged from 18 to 45 years old ($M = 22.1$, $SD = 3.39$), and 59.4% were women. All participants were instructed to visit the page of a new interactive fast-food delivery service and then to click on the "get started" button to start the interaction. Participants were then randomly assigned to the group interacting with an undisclosed or a disclosed artificial agent. In both conditions, the chat interaction was designed to support the order process by guiding users through a set of food products from which they were free to choose their option for a meal. The main questions and requests made by the chatbot were (1) to have a look at the menu, (2) to choose a dish, (3) to confirm the address or enter a new one, and (4) to enter delivery time.

Specifically, the bot started with a welcome message and guided users through the carousel menu (swiping left allowed the user to see all the pizzas displayed in the menu block). Then, the bot asked for the pizza to be ordered and directed users to click the cart symbol under the pizza image. After gathering delivery details like address and time, the bot swiftly confirmed the order, assured timely delivery and bid farewell to the user. Participants interacted with the chatbot for about five minutes. The chatbot's interaction flow is depicted in Figure 2.

In both conditions, participants experienced a chat via a messenger tool responding to the users' inputs while ordering products from the fast-food delivery service. To have a realistic approach and overcome the limitations of those studies that only use a vicarious interaction with the stimulus where users passively visualise the script of a dialogue (e.g. Chung *et al.*, 2020; Mozafari *et al.*, 2022), a fully-functional chatbot was purposely developed for the



Source(s): Authors' own work

Figure 2. Chatbot conversational flow

study. The chatbot was able to finalise the interaction with the user autonomously. The chat interface was designed with Chatfuel, an online platform that allows creating rule-based bots (https://chatfuel.com/), that is, chatbots were programmed with pre-defined rules such as pattern recognition of certain words or phrases which guaranteed a good user experience. The chat experience was created through a block builder function, meaning that users could choose from pre-defined categories by clicking on buttons. This was done to standardise the interaction and assure comparable chat interactions.

In the no-disclosure condition, participants interacted with an agent whose identity was not revealed at any time. The agent started the interaction with "*Hi {user's name}, I am here to guide you through your choice*". In the disclosed artificial identity condition, the agent was denoted as a chatbot by the chatbot's introduction *"Hi {user's name}, I am a chatbot, I am here to guide you through your choice"* at the beginning of the interaction. The sentences above were carefully curated to guarantee that they do not produce significantly different response levels across groups for variables related to the interaction partner, such as emotions, perceived friendliness, perceived expertise, perceived justice, credibility or social attractiveness.

The fast-food delivery context was chosen for the study as it represents an industry where the use of digital agents for informing about products and ordering is common (Accenture, 2021). Additionally, the service is relatively standardised among the plethora of intermediaries providing delivery services (De Cicco *et al.*, 2021), minimising the potential for brand preferences that could distort the results.

After participants were exposed to the chat stimulus, they were asked to answer a series of questions. The questionnaire administered after the treatments consisted of a first part designed to acquire demographic insights on the use of messaging apps and online purchase experience, a second part consisting of statements regarding the focal constructs, and a final part for manipulation check (asked with the dichotomous yes-no question "*Have you identified a clear and obvious disclosure about the presence of an artificial agent?*") and a question regarding previous experience with artificial sales agents.

Four participants who either failed to answer the attention checks included among the items in two scales correctly ["tick the first box (far left)"] or did not fill out the survey conscientiously or with an unrealistic completion time were discarded from further analyses. Specifically, those participants that, by checking the time log on Qualtrics, were found to have completed the survey in less than 5 minutes were excluded from the analysis since such a short completion time did not allow for a proper understanding and answering of all questions. Participants who failed the manipulation check, according to the condition they were exposed to, were also excluded from the analysis. Also, participants who completed only part of the survey by quitting before the end of the survey were excluded, which finally resulted in a total of 160 participants.

3.2 Measures, manipulation and randomisation checks

The list of the items adopted in this research can be found in Table 1. The responses were recorded on a seven-point Likert scale (1 = *"strongly disagree"*; 7 = *"strongly agree"*). Trust towards the firm was assessed according to Pengnate and Sarathy (2017) with four items; M = 5.00; SD = 1.25; Cronbach's alpha (α) = 0.86. Attitude towards the firm was measured with four items taken from Moon and Kim (2001); M = 4.82; SD = 1.42; α = 0.87. Following Bergkvist and Rossiter (2007), intention to purchase was assessed with a single item referring to the likelihood of purchasing from the firm (M = 5.11; SD = 1.75) and reflecting a behavioural expectation, which according to Sheeran (2002) possesses greater predictive validity than simple intention measures like "I intend to do X".

Table 1. Measures with literature support and scale evaluations

Constructs with items	Study 1			Study 2			Study 3		
	α	CR	AVE	α	CR	AVE	α	CR	AVE
Trust (Pengnate and Sarathy, 2017) 1. I believe that the retailer keeps its promises and commitments 2. I trust the retailer keeps customers' best interests in mind 3. The retailer is trustworthy 4. I think that the retailer will not do anything to take advantage of its customers	0.86	0.91	0.71	0.91	0.94	0.79	0.91	0.94	0.80
Attitude (Moon and Kim, 2001) 1. The retailer is good 2. The retailer is wise 3. The retailer is pleasant 4. The retailer is positive	0.88	0.92	0.73	0.90	0.93	0.73	0.94	0.96	0.85
Intention to purchase (Bergkvist and Rossiter, 2007) How likely would you purchase a product from this retailer?	–	–	–	–	–	–	–	–	–
Social presence (Gefen and Straub, 2004) 1. There is a sense of human contact in the interaction 2. There is a sense of personalness in the interaction 3. There is a sense of sociability in the interaction 4. There is a sense of human warmth in the interaction				0.93	0.94	0.79			
Speciesism 1. I think that non-human agents, such as chatbots, are in all regards inferior to humans 2. I do not mind interacting with a non-human agent when I need to interact with an organisation (r) 3. I do not think non-human agents are clever 4. I do not think non-human agents are intelligent 5. I prefer to communicate with a non-human conversational agent than interacting with a human being (r) 6. I believe that interacting with non-human agents is in all regards worse than interacting with a human 7. I cannot stand when interaction is provided through non-human agents 8. I think that non-human agents are not able to consider my unique human characteristics 9. I prefer interacting with humans rather than non-human agents							0.92	0.94	0.63

Note(s): Responses were recorded on a seven-point Likert scale. (r) = reverse-coded item
Source(s): Authors' own work

Regarding the manipulation check, in the non-human identity disclosure condition, 95% of the sample reported detecting an explicit disclosure. In the no-disclosure condition, 98.7% of participants reported not having detected a disclosure. Most of the participants (94%), regardless of the condition they were exposed to, declared to be aware of the fictitious nature of the featured fast-food delivery service which is plausible given the nature of the study as a controlled online experiment.

The two experimental groups did not statistically differ concerning gender, past experience with chatbots (which was asked at the very end of the experiment so as not to interfere with the measurement of focal constructs), messaging app usage and online purchase experience. Respondents reported daily use of messaging apps (96.9%), and 88.1% declared making online purchases between one and four times per month.

3.3 Analysis and results

A *t*-test with identity disclosure as an independent variable and trust as a dependent variable was applied to test *H1*. Trust is significantly lower in the case of non-human identity disclosure ($M = 4.81$, $SD = 1.29$) than in the no-disclosure condition ($M = 5.26$, $SD = 1.16$), $t(156) = 2.32$, $p < 0.05$, thus supporting *H1* and showing that there is a negative effect of non-human identity disclosure on trust. Means, standard deviations and *p*-values for the chatbot disclosure conditions for all studies are reported in Table 2.

The mediation effects proposed in *H2a* and *H2b* were tested by running the PROCESS model number 4 (Hayes, 2017) for each dependent variable. The manipulated independent variable was coded 1 for "no identity disclosure" and 2 for "non-human identity disclosure". In line with Hayes' (2017) recommendations, we estimated the indirect effect as well as the direct effect simultaneously using 5,000 bootstrap samples to calculate 95% bias-corrected bootstrap confidence intervals (CI). We estimated the indirect effect using the product of the coefficients approach. As reported in Table 3, the mediation analysis shows a significant negative indirect effect of identity disclosure on attitude via trust ($B = -0.16$; CI [−0.2959; −0.0196]), confirming *H2a*. We find a significant direct effect of identity disclosure on attitude ($B = -0.34$, CI [−0.5055; −0.1808]), suggesting a partial mediation through trust. The mediation analysis further reveals a significant negative indirect effect of identity disclosure on purchase intention via trust ($B = -0.13$, CI [−0.2683, −0.0186]) supporting *H2b*. Since we find no significant direct effect of non-human identity disclosure on purchase intention ($B = 0.08$, CI [−0.1742, 0.3347]), this effect is fully mediated (indirect-only). Table 3 shows the results of the mediation analysis.

4. Study 2

4.1 Design and experimental procedure

The goal of Study 2 is to examine the moderating role of social presence for the effect of the agent's non-human identity disclosure on trust towards the firm (*H3*) and subsequent outcomes (*H4a*, *H4b*). Therefore, a 2 (no identity disclosure vs non-human identity disclosure) × 2 (high vs low social presence) between-subjects design was adopted. The disclosure conditions were manipulated as in Study 1, and the same experimental design and procedure were used but enriched with a social presence condition. For the social presence condition, participants were randomly assigned to a group interacting with the chatbot set up either with low or high levels of social presence.

Following established procedures in literature (e.g. Hassanein and Head, 2007), the high social presence condition was manipulated through a socially rich interaction, which was accomplished by including paralinguistic cues in the form of socially rich picture content (GIFs and emojis), pragmatic cues (welcoming and addressing the user by name) and with

Table 2. Means, standard deviations and *p*-values for the non-human identity disclosure conditions

Study 1 Construct	Non-human identity undisclosed	Non-human identity disclosed	*p*-value
Trust	5.26 (1.16)	4.81 (1.29)	$p < 0.05$
Attitude	5.35 (1.39)	4.35 (1.39)	$p < 0.001$
Intention to purchase	5.17 (1.79)	5.06 (1.72)	$p = 0.697$

Study 2 Construct	Non-human identity undisclosed – low social presence	Non-human identity undisclosed – high social presence	Non-human identity disclosed – low social presence	Non-human identity disclosed – high social presence	*p*-value
Trust	5.35 (1.27)	5.03 (1.20)	4.23 (1.31)	5.23 (1.51)	$p < 0.001$
Attitude	5.42 (1.25)	5.39 (1.39)	4.34 (1.22)	5.32 (0.93)	$p < 0.01$
Intention to purchase	5.36 (1.43)	5.00 (1.90)	5.02 (1.51)	5.39 (1.65)	$p = 0.088$

Study 3 Construct	Non-human identity undisclosed	Non-human identity disclosed before interaction	Non-human identity disclosed within interaction	*p*-value
Trust	5.26 (0.97)	5.25 (1.03)	3.82 (1.85)	$p < 0.001$
Attitude	5.61 (1.34)	5.49 (0.94)	4.00 (2.17)	$p < 0.001$
Intention to purchase	5.32 (1.04)	5.42 (0.67)	3.94 (1.42)	$p < 0.001$

Note(s): *T*-tests were performed in Study 1, while ANOVAs were performed in Study 2 and Study 3
Source(s): Authors' own work

Table 3. Results of mediation analysis for study 1

Path	*B (SE)*	*LLCI*	*ULCI*
Direct effects			
(*H1*) Non-human identity disclosure → Trust	−0.23 (0.10) **	−0.4190	−0.0337
Non-human identity disclosure → Attitude	−0.34 (0.08) ***	−0.5055	−0.1808
Non-human identity disclosure → Intention to purchase	0.08 (0.13)	−0.1742	0.3347
Trust → Attitude	0.70 (0.07) ***	0.5650	0.8254
Trust → Intention to purchase	0.59 (0.10) ***	0.3904	0.7986
Indirect effects			
(*H2a*) Non-human identity disclosure → Trust → Attitude	−0.16 (0.07) ***	−0.2959	−0.0196
(*H2b*) Non-human identity disclosure → Trust → Intention to purchase	−0.13 (0.06) ***	−0.2683	−0.0186

Note(s): n = 160. *p < 0.05; **p < 0.01; ***p < 0.001. Number of bootstrap samples 5,000; *B* = Unstandardised coefficients (bootstrap s*tandard errors* in parentheses); LLCI = 95% lower level confidence interval; ULCI = 95% upper level confidence interval
Source(s): Authors' own work

two different registers positioned on two different points in a formal-informal language continuum. An example of the interaction with low vs high social presence is displayed in Figure 3.

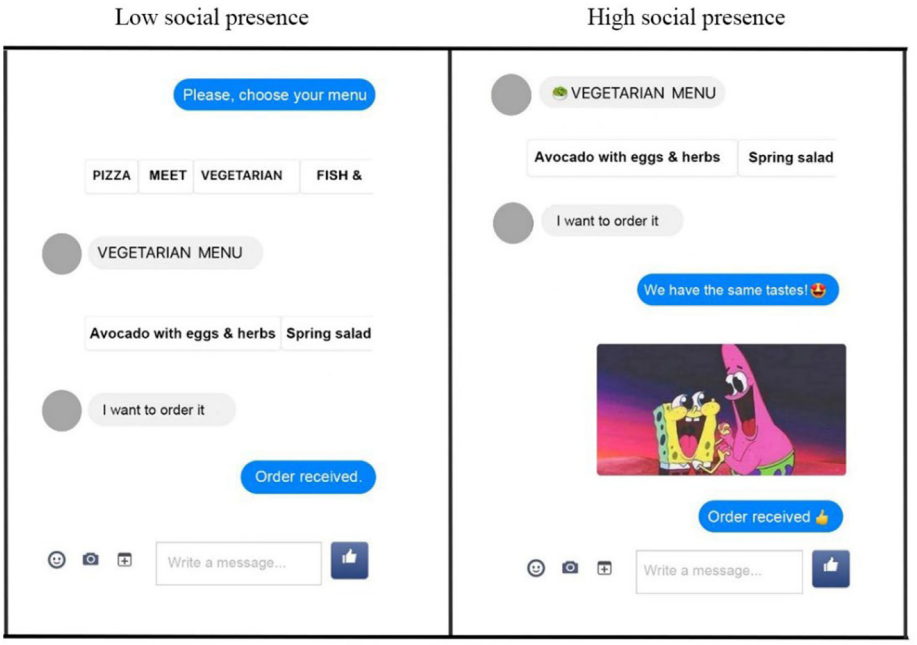
The participant recruitment process followed the same approach as in Study 1. Due to the data cleansing, which followed the same procedure reported in the method section for Study 1 (Section 3.1), nine participants were excluded from the analysis. From the final sample of 231 participants, 118 were randomly assigned to the no-disclosure condition and 113 to the disclosure condition. One hundred twenty participants interacted in the low social presence condition, and 111 in the high social presence condition. In total, there were 58 participants in the no disclosure-low social presence condition, 60 in the no disclosure-high social presence condition, 62 in the disclosure-low social presence condition and 51 in the disclosure-high social presence condition.

Participants ranged from 18 to 59 years old (M = 21.6, SD = 3.87), and 58.9% of participants were women. Almost all respondents reported daily use of messaging apps (98.7%), and 85.3% declared making online purchases between one and four times per month.

4.2 Measures, manipulation and randomisation checks

The items adopted for trust (M = 4.94; SD = 1.38; α = 0.91), attitude (M = 5.10; SD = 1.30; α = 0.90) and intention to purchase (M = 5.18; SD = 1.63) are the same as in Study 1. The manipulation for the agent disclosure was successful as 99.1% of participants in the condition with no artificial identity disclosure reported not having detected a clear disclosure about the nature of the conversational agent within the interaction. In the non-human identity disclosure condition, 98.2% correctly reported having detected a clear disclosure of the artificial nature of the agent. Most of the participants (91%), regardless of the condition they were exposed to, declared to be aware of the fictitious nature of the featured fast-food delivery service.

To check the manipulation of social presence, we asked participants to evaluate the interaction according to Gefen and Straub's (2004) social presence scale with four items (e.g. "There is a sense of human contact in the interaction", see Table 1), M = 4.20; SD = 1.68; α = 0.93. The analysis showed that perceived social presence is significantly lower in the low social presence condition (M = 3.27, SD = 1.44) compared to the high social presence condition (M = 5.21, SD = 1.31); $F(1, 230) = 113.47$, $p < 0.001$, $\eta^2 = 0.33$, thus suggesting



Note(s): High social presence was achieved through GIFs, emoticons and socially rich text
Source(s): Authors' own work

Figure 3. Examples of chatbot interactions with low vs high social presence

that the manipulation of social presence was successful. We also checked that the social presence perceived was not statistically different between the two disclosure conditions [no disclosure: $M = 4.69$, $SD = 1.87$; disclosure: $M = 4.33$, $SD = 1.99$; $F(1, 230) = 2.08$, $p = 0.15$]. The four experimental groups did not statistically differ concerning gender, age and past experience with chatbots.

4.3 Analysis and results

4.3.1 Interaction effects of non-human identity disclosure and social presence on trust towards the firm. The main effects of non-human identity disclosure confirm the results of Study 1: the disclosure has a significant impact on trust [$F(1, 230) = 6.79$; $p < 0.01$; partial $\eta^2 = 0.03$]. In case the disclosure is present, trust decreases ($M = 4.68$, $SD = 1.49$) compared to the no-disclosure condition ($M = 5.19$, $SD = 1.24$).

An ANOVA was then applied to test *H3* and compare the interaction effects of identity disclosure and social presence (high vs low) on trust. The two-way interaction effect between disclosure and social presence on trust is significant [$F(1,230) = 14.13$, $p < 0.001$; $\eta^2 = 0.05$]. Thus, *H3* is supported. We conducted a simple main effect analysis to further disentangle this cross-over interaction. The results show that the effect of identity disclosure is not significant in the high social presence condition [no disclosure: $M = 5.03$, $SD = 1.20$; disclosure: $M = 5.23$, $SD = 1.51$; $F(1, 230) = 0.63$; $p = 0.42$]. However, in the low social presence condition, we find a highly significant negative effect of identity disclosure on trust [$F(1, 230) = 21.14$;

$p < 0.001$; partial $\eta^2 = 0.85$]. In the case of low social presence, trust is higher when the identity disclosure is absent ($M = 5.35$, $SD = 1.27$) compared to when disclosure is present ($M = 4.23$, $SD = 1.31$). This suggests that in the case of low social presence, potential buyers place greater trust in the firm when non-human cues are prevented. In case the non-human identity disclosure is present, trust decreases ($M = 4.68$, $SD = 1.49$), compared to when the non-human identity disclosure is absent ($M = 5.19$, $SD = 1.24$). We also checked for possible effects of social presence on trust: the direct effect of social presence on trust is not significant (low social presence: $M = 4.77$, $SD = 1.40$; high social presence: $M = 5.12$, $SD = 1.35$), $F(1, 230) = 3.71$, $p = 0.07$; partial $\eta^2 = 0.01$. Given these results, social presence can be characterised as a pure moderator (Sharma *et al.*, 1981).

4.3.2 Moderated mediation effects. To test *H4a* and *H4b*, we used PROCESS model number 7 for moderated mediation with 5,000 bootstrap samples with identity disclosure as the independent variable, coded 1 for "no identity disclosure" and 2 for "non-human identity disclosure", trust as the mediator, and social presence as the moderator. The results show that the mediated effects of non-human identity disclosure on attitude and purchase intention via trust are moderated by social presence. The significant interaction between the moderator variable (social presence) and the predictor variable (identity disclosure) [$F(1, 230) = 14.13$, $p < 0.001$], and the non-significant effect of the moderator on the dependent variable (trust) [$F(1, 230) = 3.71$, $p = 0.07$] confirm the nature of social presence as a "pure moderator" according to the Sharma *et al.* (1981) terminology. For attitude, as expected, the indirect effect of the interaction between the artificial identity disclosure and social presence on attitude via trust is significant [$B = 0.38$, CI (0.1620, 0.6240)]; thus, *H4a* is supported. We find no direct effect of the interaction of identity disclosure and social presence on attitude [$B = 0.07$, CI (−0.0591, 0.2050)]. Hence, a full moderated mediation exists between identity disclosure and attitude. In line with *H3*, the indirect effect of the disclosure on attitude via trust is significant in the low social presence condition [$B = -0.31$, CI (−0.4674, −0.1790)].

For intention to purchase, the indirect effect of the interaction between identity disclosure and social presence on purchase intention via trust is significant ($B = 0.31$, CI [0.1219, 0.5368]), thus supporting *H4b*. We find no direct effect of the interaction between identity disclosure and social presence on purchase intention ($B = 0.03$, CI [−0.1734, 0.2323]). Hence, there is a full moderated mediation between the identity disclosure and purchase intention. Also, in line with *H3*, the indirect effect via trust is significant only in the low social presence condition ($B = -0.26$, CI [−0.4136, −0.1371]). Results for the direct and indirect effects are reported in Table 4.

5. Study 3

5.1 Design and experimental procedure

The goal of Study 3 is to examine the moderating role of speciesism in relation to the effects of disclosing the non-human identity of the agent at different stages of the purchase journey (before the interaction vs within the interaction) on trust towards the firm (*H5*) and subsequent outcomes (*H6a*, *H6b*). Although we expect that speciesism against artificial agents should not play a role in a no-disclosure condition due to a lack of any machine-related cues, we include a no-disclosure condition also in Study 3 to replicate the test of *H1* by using a more fine-grained scheme of disclosure strategies (disclosure before vs within the interaction), testing the robustness of the trust-declining effect of disclosure. Therefore, a 3 (no disclosure vs non-human identity disclosure before the interaction vs non-human identity disclosure within the

Table 4. Results of moderation and moderated mediation analysis for Study 2

Path	*B (SE)*	*LLCI*	*ULCI*
Direct effects			
Non-human identity disclosure → Trust	−0.23 (0.09) **	−0.3995	−0.0555
Non-human identity disclosure → Attitude	−0.17 (0.07) *	−0.3000	−0.0381
Non-human identity disclosure → Intention to purchase	0.12 (0.10)	−0.0771	0.3209
Trust → Attitude	0.57 (0.05) ***	0.4775	0.6667
Trust → Intention to purchase	0.47 (0.07) ***	0.3275	0.6150
Social presence → Trust	0.17 (0.09)	−0.0039	0.3401
Social presence → Attitude	0.14 (0.07) *	0.0143	0.2725
Social presence → Intention to purchase	−0.08 (0.10)	−0.2744	0.1224
Moderated effect			
(*H3*) Non-human identity disclosure x social presence → Trust	0.33 (0.09) ***	0.1561	0.5001
Moderated mediation effects			
(*H4a*) Non-human identity disclosure x social presence → Trust → Attitude	0.38 (0.12) ***	0.1620	0.6240
(*H4b*) Non-human identity disclosure x social presence → Trust → Intention to purchase	0.31 (0.11) ***	0.1219	0.5368
Conditional indirect effects of non-human identity disclosure on attitude			
Low social presence	−0.31 (0.07) ***	−0.4674	−0.1790
High social presence	0.06 (0.08)	−0.0816	0.2256
Conditional indirect effects of non-human identity disclosure on intention to purchase			
Low social presence	−0.26 (0.07) ***	−0.4136	−0.1371
High social presence	0.05 (0.07)	−0.0684	0.1893

Note(s): n = 231. *p < 0.05; **p < 0.01; ***p < 0.001. Number of bootstrap samples 5,000; B = Unstandardised coefficients (bootstrap *standard errors* in parentheses); LLCI = 95% lower level confidence interval; ULCI = 95% upper level confidence interval
Source(s): Authors' own work

interaction) × speciesism (continuous) between-subjects design was adopted. To ensure consistency across studies, we selected the same context (fast-food delivery service) and used the same experimental design and procedure as in Study 1, with the addition of a further disclosure condition. As in the previous studies, in the no-disclosure condition, there was no mention of the artificial identity of the agent at any point. In the before-interaction disclosure condition, the participant viewed a page with the statement "*You are going to interact with a chatbot supported by artificial intelligence*" before starting the interaction with the chatbot. In the within-interaction disclosure condition, the user was informed about the artificial identity of the agent during the interaction (i.e. "*I am a programmed textbot*", "*Thank you for ordering with the chatbot Pizzabot!*").

Data cleansing followed the same procedure reported in the method section for Studies 1 and 2 and resulted in 39 responses excluded. The final sample included 214 participants ranging from 18 to 61 years old (M = 34.7, SD = 10.75), and 46.7% were women. Seventy-one participants were randomly assigned to the no-disclosure condition, 72 participants to the before-interaction disclosure condition and 71 to the within-interaction disclosure condition. Most respondents reported daily use of messaging apps (86.9%), and 82.2% declared making online purchases between one and four times per month.

5.2 Measures, manipulation and randomisation checks

The items adopted for trust (M = 4.78; SD =1.50; α = 0.91), attitude (M = 4.90; SD = 1.42; α = 0.94) and intention to purchase (M = 5.03; SD = 1.72) are the same as in Studies 1 and 2. Speciesism, our measured moderator, was captured through a purposely developed speciesism scale (M = 3.91; SD = 1.43; α = 0.92.), including nine items reported in Table 1. We explain the scale adaptation process in the following sections.

Speciesism is an established concept in biology and philosophy, rooted in the work of Ryder (2006) and Singer and Mason (2007). It refers to a sense of superiority that humans hold over animals. Caviola *et al.* (2019) put forward a first measure that transfers the concept to the field of social psychology. Building on these foundations, Schmitt (2020) further extended the concept of speciesism to encompass non-human entities more broadly (i.e. transcending animals) by theorising how people may perceive themselves as superior to artificial agents. However, there is no scale for measuring speciesism in terms of beliefs of superiority displayed by humans towards artificial agents. Thus, we adapted the 13 item-scale of Caviola *et al.* (2019) to the context of interactions with artificial agents, following a multi-stage pre-study. In the first stage, we discussed these items with a group of experts in the disciplines of IT, marketing, psychology and behavioural economics. These expert discussions revealed that items oriented towards sexual and moral harassment covered in the original item list of Caviola *et al.* (2019) were deemed non-applicable to artificial agents. The same applied to items reflecting property rights and the possibility to do whatever individuals want with non-human entities (like animals) for their own pleasure. Hence, these items were omitted as they were better suited for capturing speciesism towards animals.

In the second stage, we discussed the remaining items with a group of regular consumers who had not yet acquired specific knowledge or interest in artificial agents. A total of 12 consumers were assigned to two focus groups. The interviews were balanced in terms of sample representativity (age, gender, and educational background) and lasted for approximately one hour, moderated by one of the authors. The expert panel reviewed the outcomes from the focus groups. During the interviews, participants frequently emphasised aspects concerning the inferiority of artificial agents like chatbots relative to humans as typical for speciesism-related prejudice. This was associated with chatbot characteristics such as naivety, stupidity, a lack of fine-grained intelligence, an inability to adapt to the flow of a conversation, reinforcement of trivial issues and an inability to use or understand second meanings. Hence, individual items were added.

The two qualitative stages finally resulted in nine items that refer to the superiority of humans and/or the inferiority of artificial agents in service encounters. All these items reflect prejudices that are applied to evaluating the idiosyncrasies of interactions between humans and machines. In a third stage, we subjected these items to a quantitative pre-test (n = 41, 53.66% women, mean age: 32.36) that resulted in adequate validity and reliability of the scale with all fit measures ranging above the recommended thresholds (α = 0.87; construct reliability = 0.92; AVE = 0.57).

The manipulation check for the experimental conditions was conducted by asking participants whether they had perceived a clear disclosure about the nature of the conversational agent before the start of or within the interaction, respectively. In the no-disclosure condition, 98.6% of participants reported not detecting a disclosure. In the before-interaction disclosure condition, 95.8% of participants reported having detected an explicit disclosure before the interaction. In comparison, 98.6% of participants declared having seen a clear disclosure during the interaction in the within-interaction disclosure condition. The experimental groups did not show statistically significant differences in terms of gender, age and experience with chatbots.

5.3 Analysis and results

5.3.1 Interaction effects of non-human identity disclosure and speciesism on trust towards the firm. In Study 3, we test how consumers respond to different disclosure strategies depending on the speciesism stimulation. In *H5*, we hypothesised a negative moderating effect of speciesism on the relationship between non-human identity disclosure before the interaction (vs within the interaction) and trust. We relied on Hayes' PROCESS model number 1 to perform the analysis, including the three identity disclosure conditions of our categorical independent variable, and speciesism as moderator. As our moderator is a continuous variable, to test conditional effects, we followed Hayes' (2017) approach as well as other recent studies (e.g. Argouslidis *et al.*, 2018) and conducted a median split of the sample (*median* = 3.77).

The overall equation was significant: $R^2 = 0.52$, $F(3, 210) = 75.31$, $p < 0.001$. Congruent with *H5*, we find a significant negative effect of the interaction between non-human identity disclosure conditions and speciesism on trust [$B = -0.12$, CI (−0.2035; −0.0334)]. The effect of disclosure on trust was not statistically significant at the low speciesism level, but it was significant for the high speciesism group. Specifically, we found that the relationship between identity disclosure conditions and trust was negative and significant among people with a high level of speciesism [$B = -0.42$, CI (−0.5817; −0.2610)]. In contrast, the relationship between identity disclosure and trust was insignificant for low speciesism [$B = 0.07$, CI (−0.2604; 0.1178)]. In this case, according to the framework proposed by Sharma *et al.* (1981), given the significant interaction between our moderator (speciesism) and the predictor variable (identity disclosure) [$F(3, 210) = 75.31$, $p < 0.001$], and the significant relation between the moderator and the dependent variable (trust) [$B = -0.30$, CI (−0.5575; −0.0428), $p < 0.01$], we can classify speciesism as a "quasi moderator" and suggest that it functions more as an independent variable than as a traditional moderator.

The results provide interesting insights demonstrating that priming non-human identity cues reduces trust when it happens within the chatbot interaction (vs before). This negative impact is observed exclusively for individuals with high speciesism, while for people low in speciesism, there is no decline in trust. Hence, for individuals low in speciesism, no matter at which stage of the purchase journey the non-human identity is disclosed, there is no negative effect on trust.

Furthermore, in replicating the results for testing *H1*, when presenting a non-human identity disclosure within the chatbot interaction, trust is significantly lower than in the case of no disclosure ($M_{within} = 3.82$, $SD = 1.85$; $M_{no} = 5.26$, $SD = 0.97$; $p < 0.001$). However, strikingly, there is no significant trust difference between before-interaction disclosure and no disclosure ($M_{before} = 5.25$, $SD = 1.03$; $M_{no} = 5.26$, $SD = 0.97$; $p = 0.97$).

5.3.2 Moderated mediation effects. To test *H6*, as in Study 2, we used PROCESS model number 7 for moderated mediation with 5,000 bootstrap samples with the three identity disclosure conditions as the independent variable, trust as the mediator and speciesism as the moderator. As in the previous moderation analysis, since our moderator is a continuous variable, to test conditional effects, we split the sample based on the median, following Hayes' (2017) approach, as well as other recent studies (e.g. Argouslidis *et al.*, 2018).

The results indicate that the effects of identity disclosure on attitude and intention to purchase via trust are moderated by speciesism. For attitude, the negative indirect effect of the interaction between identity disclosure and speciesism on attitude via trust is significant ($B = -0.09$, CI [−0.1700, −0.0050]). The indirect effect of identity disclosure on attitude via trust is significant and negative for high speciesism ($B = -0.31$, CI [−0.4824, −0.1553]). Hence, *H6a* is supported. Additionally, for intention to purchase, the indirect effect of the interaction between identity disclosure and speciesism on purchase intention via trust is

negative and significant (B = −0.09, CI [−0.1785, −0.0068]). The indirect effect of the disclosure on intention to purchase via trust is negatively significant for high speciesism (B = −0.33, CI [−0.5055, −0.1558]). Therefore, *H6b* is also supported. Results for the direct and indirect effects are reported in Table 5.

6. General discussion

Digital agents are capable of simulating human conversations so convincingly that consumers may have difficulty distinguishing them from human assistants (Schmitt, 2020). This can lead to perceptions of deception, making identity disclosure a recommended practice to ensure a more balanced interaction between interlocutors (Murtarelli *et al.*, 2021). However, the circumstances under which signalling the non-human nature of these agents might have negative effects, as well as the strategies to prevent such drawbacks, remain largely unexplored. Our research explores this timely topic and provides new insights into how disclosing a sales agent's artificial identity impacts trust in the firm, consumer attitudes and purchase outcomes. We also examine how disclosure can be managed to minimise trust declines and mitigate negative consumer responses, while considering the role of consumer speciesism in shaping reactions to disclosure.

Study 1 reveals that when non-human cues are introduced in an interaction, people tend to discriminate against artificial agents, undermining trust and indirectly reducing purchase

Table 5. Results of moderation and moderated mediation analysis for study 3

Path	*B (SE)*	*LLCI*	*ULCI*
Direct effects			
Non-human identity disclosure → Trust	0.22 (0.19)	−0.1555	0.5918
Non-human identity disclosure → Attitude	−0.11 (0.05) **	−0.2090	−0.0211
Non-human identity disclosure → Intention to purchase	−0.17 (0.07) ***	−0.3091	−0.0256
Trust → Attitude	0.75 (0.04) ***	0.6697	0.8258
Trust → Intention to purchase	0.77 (0.06)	−0.6540	0.8897
Speciesism → Trust	−0.30 (0.13) **	−0.5575	−0.0428
Speciesism → Attitude	−0.17 (0.12)	−0.4074	0.0504
Speciesism → Intention to purchase	−0.11 (0.15) ***	−3.111	−0.1125
Moderated effect			
(*H5*) Non-human identity disclosure x speciesism → Trust	−0.12 (0.04) **	−0.2035	−0.0334
Moderated mediation effects			
(*H6a*) Non-human identity disclosure x speciesism → Trust → Attitude	−0.09 (0.04) ***	−0.1700	−0.0050
(*H6b*) Non-human identity disclosure x speciesism → Trust → Intention to purchase	−0.09 (0.04) ***	−0.1785	−0.0068
Conditional indirect effects of non-human identity disclosure on attitude			
Low speciesism	−0.05 (0.08)	−0.2267	0.1081
High speciesism	−0.31 (0.08) ***	−0.4824	−0.1553
Conditional indirect effects of non-human identity disclosure on intention to purchase			
Low speciesism	−0.05 (0.08)	−0.2248	0.1099
High speciesism	−0.33 (0.09) ***	−0.5055	−0.1558

Note(s): n = 214. *p < 0.05; **p < 0.01; ***p < 0.001. Number of bootstrap samples 5,000; B = Unstandardised coefficients (bootstrap *standard errros* in parentheses); LLCI = 95% lower level confidence interval; ULCI = 95% upper level confidence interval
Source(s): Authors' own work

intention, even when the artificial assistant functions as an optimal sales agent. The results are in line with research emphasising the negative effects of conversing with chatbots compared to humans in terms of lower user acceptance (Murgia *et al.*, 2016) and more negative emotional reactions (Ciechanowski *et al.*, 2019).

Study 2 demonstrates that social presence is crucial for attenuating resistance to artificial agents, but only at certain levels. When the agent's social presence is low, disclosure translates into lower perceived trust, and, in turn, lower attitude and intention to purchase. Probably, at a low level of social presence, users struggle to perceive empathy and emotional resonance, which in turn affects consumer responses (Chen *et al.*, 2021). Interestingly, there is no effect in the case of high social presence, indicating that once a certain level of social presence is reached, this no longer contributes to strengthening trust and subsequent responses to artificial identity disclosure.

Study 3 reveals that consumers with higher speciesism exhibit lower trust and less favourable responses than those with lower prejudice when non-human cues are presented during the interaction. However, these deleterious effects are not observed when the cues are introduced before the interaction. Disclosing the non-human identity during the interaction can interrupt the conversation and serve as an intrusive reminder of the artificial nature of the agent. This reminder may activate resistance in consumers, even leading to a direct negative impact on purchase intention. Users may "fool" themselves by believing they are engaged in a human-like interaction and this flow might be disrupted by non-human cues from the agent. The results support Mende *et al.*'s (2019) view that highlighting an agent's machine-like nature is not always problematic (in our case, when it happens at the very beginning of the journey before the interaction begins).

6.1 Theoretical implications

Our findings make important contributions to the marketing literature regarding how far consumers accept artificial sales agents. Our results contribute to answering the recent calls to understand better the determinants of successful human-machine integration (Chandra *et al.*, 2022; Pizzi *et al.*, 2021), and lay possible foundations for the game-changing role of Generative AI-based interactions (Peres *et al.*, 2023) which will probably be the new normal in conversational commerce.

First, we show that employing digital agents and explicitly labelling them as such can be a double-edged sword for companies. While transparency reduces information asymmetry and mitigates risk and privacy concerns (Silva *et al.*, 2023; Murtarelli *et al.*, 2021), revealing an agent's artificial identity creates a negative touchpoint by eroding rather than enhancing users' trust.

Second, by including considerations from speciesism and measuring it through a newly developed scale, we are the first to respond to Schmitt's (2020) call for empirical research on how speciesism against non-human entities unfolds in technology-related contexts, and specifically on the role of such speciesism bias for shaping trust. Our speciesism scale enables the segmentation of individuals based on their level of speciesism through a median split and demonstrates strong validity and reliability, indicating significant potential for future applications. By advancing Luo *et al.*'s (2019) findings on voice bots, we demonstrate that priming users with information about the agent's artificial identity makes speciesism bias salient, negatively shaping the relationship between the use of text-based sales assistants and relevant business outcomes. This confirms speciesism as a discriminatory mechanism driving trust erosion after disclosure, with crucial implications: highly "speciesist" individuals extend their prejudice against machines to firm evaluations, even in the absence of actual performance differences. Our findings theoretically contribute to the human-AI

interaction literature researching negative reactions towards chatbots which are not guided by their objective performance but by mental biases (Schmitt, 2020). We believe our results could be the starting point to investigate the causes of this prejudice and to detect design elements and interaction styles that help to fight speciesism. In this vein, although speciesism was initially conceptualised and examined as a moderator in our study, the results of Study 3 indicate that speciesism also emerged as a significant driver of downstream outcomes. This finding adds to the literature on speciesism against artificial agents, demonstrating that speciesism independently influences consumer trust and intentional responses, suggesting that pre-existing biases against non-human entities may directly shape consumers' perceptions. This insight underscores the need for further investigation into speciesism as a stand-alone factor impacting consumer-agent interactions.

Thirdly, this work enriches social presence theory (Short *et al.*, 1976) by situating it within interactions involving artificial agents, where social presence serves as a moderating factor that shapes consumer perceptions. We show that social presence plays a critical role in improving consumer attitudes while minimising friction in chatbot-mediated interactions and transforming potential drawbacks into opportunities. In doing so, the study offers evidence for the adaptability of social presence theory to technologically advanced, AI-driven environments.

6.2 Managerial implications

Although artificial sales agents can offer several benefits for firms throughout the customer journey, such as collecting information, providing better support in the shopping process and promoting the brand and the company (Roy and Naidoo, 2021), we want to alert managers that employing such conversational agents can also present challenges related to disclosing their artificial nature.

As good news for firms, our results show that negative reactions stemming from the explicit disclosure of the artificial identity, which is advisable for firms to ensure transparency, can be mitigated when combined with specific social cues. If managers design the interface of human-agent interactions in a sociable way, they can break down the barrier between humans and machines. Properly implementing features of digital agents in a way that signals empathy and emotional capacity enhances value exchange (Taylor *et al.*, 2020). Our findings help marketers reduce consumer resistance to artificial agents by applying specific interaction cues that enhance social presence and bridge the gap between humans and machines, strengthening trust, which might otherwise decline when consumers are explicitly aware of the agent's artificial identity.

Our insights provide strategies for firms to manage transparency constructively, helping them handle a disclosure while minimising possible adverse effects of being open about the machine identity of sales agents. One way to achieve this is to disclose the artificial nature of a sales agent as early as possible in the customer journey, preferably before the conversation begins. Our results show that reminding consumers during the interaction that they are conversing with an artificial entity activates speciesism-related prejudices, leading to negative reactions towards artificial employees. Obviously, such identity priming interrupts the smooth and easy-going interaction and might hinder a "mindless anthropomorphisation" of the digital partner from unfolding (Kim and Sundar, 2012). If revealing non-human cues within an interaction is inevitable, managers can use social mechanisms when presenting the agent within the interaction to counteract negative perceptions associated with artificial sales agents.

Companies are increasingly relying on chatbots for a vast array of tasks. At the same time, we observe the moderating effect of speciesism, which requires managers to properly select the audience they target with chatbots, to minimise the risk of undesirable effects. Speciesism will not easily fade away, resulting from a biological-evolutionary process

deeply ingrained in humans (Schmitt, 2020). Therefore, more thought should be given to how chatbots are designed. We recommend that firms design their digital bots not to activate speciesism, as this can lead to an undesirable decline in trust. Characteristics that normally activate speciesism (naivety, failure in maintaining a human-like interaction, repetitiveness, definition-insisting) should be avoided. Companies willing to employ artificial agents in their marketing strategy should target consumers with low speciesism as they will not react more unfavourably to artificial entities than they would react to human employees. Alternatively, in cases where individuals exhibit a higher level of speciesism, the need for a well-designed bot is even more crucial as it needs to be designed in a way that reduces any apparent bias, facilitating smoother interactions.

6.3 Limitations and future research

Despite the advances, our research has limitations that warrant further investigation into the evolving consumer behaviour towards chatbots. First, although the study used a real chatbot and overcame the limitations associated with presenting participants with mere screenshots of chat interactions, participants engaged in a simulated purchase task. Future studies could explore chatbot interactions with real-life firms to enhance external validity. Furthermore, as our study participants were predominantly young, further research could consider different age groups, as key constructs in our study, such as speciesism and trust, may vary by age. Similarly, the same business category was used in all three studies to allow a robust comparison of results across studies. Future studies should consider different service categories or products with varied purchasing values to enhance generalisability.

Second, our model investigates a limited number of variables or settings. Future research could explore the role of speciesism in relation to different performance levels and even chatbot failures, building on Mozafari *et al's.* (2022) work or examine Longoni and Cian's (2022) framework on opposing attribute trade-offs (e.g. hedonic vs utilitarian). It would also be interesting to investigate whether personality characteristics, such as the dispositional inclination to resist change, the need for human interaction or the cultural background of participants, could moderate the effects of identity disclosure. Additionally, research could also explore whether individuals distrust algorithmic advice more than human advice, focusing specifically on chatbots applied in conversational commerce.

Third, we relied on linguistic and paralinguistic cues concerning social presence. Further studies should explore the impact of a broader range of social aspects that might work as moderators of the relationship between disclosure and trust, such as local habits for greeting people (e.g. through geo-referencing), (regional) jargon or customised humour. It would also be interesting to verify whether these findings could be transferred to the universe of voice assistants, where the services' intangibility causes greater uncertainty for consumers (e.g. when providing financial advice). As speciesism against artificial agents is not going to be entirely eradicated through high social presence or the right timing of the agent's introduction alone, and as we will soon be called to interact daily with AI-leveraged conversational technologies, other ways to improve feelings towards these agents are welcome, especially for people displaying high-speciesism levels.

Finally, although we believe that firms need to understand possible oppositions towards artificial agents, we equally agree that it is challenging to administer a 9-item measure to discern if a user scores low or high in speciesism. In cases where simplicity is key, we suggest that firms use a shorter version or even a single item from our speciesism scale, administered before the interaction with the chatbot, to offer specific conversational paths contingent on the response to the question(s). Based on the indicator reliabilities and item-to-total correlations, the three strongest indicators of speciesism suitable for a short scale are: "*I*

believe that interacting with non-human agents is in all regards worse than interacting with a human"; "*I think that non-human agents are not able to consider my unique human characteristics*"; and "*I prefer interacting with humans rather than non-human agents*". Alternatively, researchers could infer such individual disposition through a sentiment analysis based on preliminary interactions with the agent.

References

Accenture (2021), "Food shopping gets a dash of innovation", available at: www.accenture.com/lu-en/case-studies/retail/coop-italia-reinvents-grocery-shopping (accessed 6 November 2022).

Argouslidis, P., Skarmeas, D., Kühn, A. and Mavrommatis, A. (2018), "Consumers' reactions to variety reduction in grocery stores: a freedom of choice perspective", *European Journal of Marketing*, Vol. 52 Nos 9/10, pp. 1931-1955.

Aydin, S. and Özer, G. (2005), "The analysis of antecedents of customer loyalty in the Turkish mobile telecommunication market", *European Journal of Marketing*, Vol. 39 Nos 7/8, pp. 910-925.

Bargh, J.A., Chen, M. and Burrows, L. (1996), "Automaticity of social behavior: direct effects of trait construct and stereotype activation on action", *Journal of Personality and Social Psychology*, Vol. 71 No. 2, pp. 230-244.

Bergkvist, L. and Rossiter, J.R. (2007), "The predictive validity of multiple-item versus single-item measures of the same constructs", *Journal of Marketing Research*, Vol. 44 No. 2, pp. 175-184.

Blut, M., Wang, C., Wünderlich, N.V. and Brock, C. (2021), "Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI", *Journal of the Academy of Marketing Science*, Vol. 49 No. 4, pp. 632-658.

Bozkurt, S., Gligor, D.M. and Babin, B.J. (2020), "The role of perceived firm social media interactivity in facilitating customer engagement behaviors", *European Journal of Marketing*, Vol. 55 No. 4, pp. 995-1022.

Castaldo, S., Premazzi, K. and Zerbini, F. (2010), "The meaning(s) of trust: a content analysis on the diverse conceptualizations of trust in scholarly research on business relationships", *Journal of Business Ethics*, Vol. 96 No. 4, pp. 657-668.

Castelo, N., Bos, M.W. and Lehman, D.R. (2019), "Task-Dependent algorithm aversion", *Journal of Marketing Research*, Vol. 56 No. 5, pp. 809-825.

Castelo, N., Boegershausen, J., Hildebrand, C. and Henkel, A.P. (2023), "Understanding and improving consumer reactions to service bots", *Journal of Consumer Research*, Vol. 50 No. 4, pp. 848-863.

Caviola, L., Everett, J.A.C. and Faber, N.S. (2019), "The moral standing of animals: towards a psychology of speciesism", *Journal of Personality and Social Psychology*, Vol. 116 No. 6, pp. 1011-1029.

Chandra, S., Shirish, A. and Srivastava, S.C. (2022), "To be or not to be… human? Theorizing the role of human-like competencies in conversational artificial intelligence agents", *Journal of Management Information Systems*, Vol. 39 No. 4, pp. 969-1005.

Chen, N., Mohanty, S., Jiao, J. and Fan, X. (2021), "To err is human: tolerate humans instead of machines in service failure", *Journal of Retailing and Consumer Services*, Vol. 59 No. 2, p. 102363.

Chung, M., Ko, E., Joung, H. and Kim, S.J. (2020), "Chatbot e-service and customer satisfaction regarding luxury brands", *Journal of Business Research*, Vol. 117, pp. 587-595.

Ciechanowski, L., Przegalinska, A., Magnuski, M. and Gloor, P. (2019), "In the shades of the uncanny valley: an experimental study of human–chatbot interaction", *Future Generation Computer Systems*, Vol. 92, pp. 539-548.

Davenport, T., Guha, A., Grewal, D. and Bressgott, T. (2020), "How artificial intelligence will change the future of marketing", *Journal of the Academy of Marketing Science*, Vol. 48 No. 1, pp. 24-42.

De Cicco, R., Silva, S.C. and Alparone, F.R. (2020), "Millennials' attitude toward chatbots: an experimental study in a social relationship perspective", *International Journal of Retail and Distribution Management*, Vol. 48 No. 11, pp. 1213-1233.

De Cicco, R., Silva, S.C. and Alparone, F.R. (2021), "'It's on its way': chatbots applied for online food delivery services, social or task-oriented interaction style?", *Journal of Foodservice Business Research*, Vol. 24 No. 2, pp. 140-164.

Dietvorst, B.J., Simmons, J.P. and Massey, C. (2015), "Algorithm aversion: people erroneously avoid algorithms after seeing them err", *Journal of Experimental Psychology: General*, Vol. 144 No. 1, pp. 114-126.

Elmashhara, M.G., De Cicco, R., Silva, S.C., Hammerschmidt, M. and Silva, M.L. (2024), "How gamifying AI shapes customer motivation, engagement, and purchase behavior", *Psychology and Marketing*, Vol. 41 No. 1, pp. 134-150.

Evans, N.J., Phua, J., Lim, J. and Jun, H. (2017), "Disclosing Instagram influencer advertising: the effects of disclosure language on advertising recognition, attitudes, and behavioral intent", *Journal of Interactive Advertising*, Vol. 17 No. 2, pp. 138-149.

Fishbein, M. and Ajzen, I. (1977), "Belief, attitude, intention, and behavior: an introduction to theory and research", *Philosophy and Rhetoric*, Vol. 10 No. 2, pp. 130-132.

Følstad, A., Nordheim, C.B. and Bjørkli, C.A. (2018), "What makes users trust a chatbot for customer service? An exploratory interview study", *Proceedings of the International Conference on Internet Science, INSCI 2018*, Vol. 5, pp. 194-208.

Fuegener, A., Grahl, J., Gupta, A. and Ketter, W. (2022), "Cognitive challenges in human–artificial intelligence collaboration: investigating the path toward productive delegation", *Information Systems Research*, Vol. 33 No. 2, pp. 678-696.

Garvey, A.M., Kim, T. and Duhachek, A. (2022), "Bad news? Send an AI. Good news? Send a human", *Journal of Marketing*, Vol. 87 No. 1, pp. 10-25.

Gefen, D. and Straub, D.W. (2004), "Consumer trust in B2C e-commerce and the importance of social presence: experiments in e-products and e-services", *Omega*, Vol. 32 No. 6, pp. 407-424.

Gu, H., Zhang, T., Lu, C. and Song, X. (2021), "Assessing trust and risk perceptions in the sharing economy: an empirical study", *Journal of Management Studies*, Vol. 58 No. 4, pp. 1002-1032.

Hassanein, K. and Head, M. (2007), "Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping", *International Journal of Human-Computer Studies*, Vol. 65 No. 8, pp. 689-708.

Hayes, A.F. (2017), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, Guilford Publications, New York City, NY.

Holthöwer, J. and Van Doorn, J. (2023), "Robots do not judge: service robots can alleviate embarrassment in service encounters", *Journal of the Academy of Marketing Science*, Vol. 51 No. 4, pp. 767-784.

Hong, I.B. and Cha, H.S. (2013), "The mediating role of consumer trust in an online merchant in predicting purchase intention", *International Journal of Information Management*, Vol. 33 No. 6, pp. 927-939.

Huang, M.H. and Rust, R.T. (2018), "Artificial intelligence in service", *Journal of Service Research*, Vol. 21 No. 2, pp. 155-172.

Iacobucci, S., De Cicco, R., Michetti, F., Palumbo, R. and Pagliaro, S. (2021), "Deepfakes unmasked: the effects of information priming and bullshit receptivity on deepfake recognition and sharing intention", *Cyberpsychology, Behavior, and Social Networking*, Vol. 24 No. 3, pp. 194-202.

Kaplan, A. and Haenlein, M. (2020), "Rulers of the world, unite! The challenges and opportunities of artificial intelligence", *Business Horizons*, Vol. 63 No. 1, pp. 37-50.

Kim, Y. and Sundar, S.S. (2012), "Anthropomorphism of computers: is it mindful or mindless?", *Computers in Human Behavior*, Vol. 28 No. 1, pp. 241-250.

Kim, Y. and Peterson, R.A. (2017), "A meta-analysis of online trust relationships in E-commerce", *Journal of Interactive Marketing*, Vol. 38 No. 1, pp. 44-54.

Kopalle, P.K., Gangwar, M., Kaplan, A., Ramachandran, D., Reinartz, W. and Rindfleisch, A. (2021), "Examining artificial intelligence (AI) technologies in marketing via a global lens: current trends and future research opportunities", *International Journal of Research in Marketing*, Vol. 39 No. 2, pp. 522-540.

Li, X., Hess, T.J. and Valacich, J.S. (2008), "Why do we trust new technology? A study of initial trust formation with organizational information systems", *The Journal of Strategic Information Systems*, Vol. 17 No. 1, pp. 39-71.

Liu-Thompkins, Y., Okazaki, S. and Li, H. (2022), "Artificial empathy in marketing interactions: bridging the human-AI gap in affective and social customer experience", *Journal of the Academy of Marketing Science*, Vol. 50 No. 6, pp. 1198-1218.

Longoni, C. and Cian, L. (2022), "Artificial intelligence in utilitarian vs. hedonic contexts: the 'word-of-machine' effect", *Journal of Marketing*, Vol. 86 No. 1, pp. 91-108.

Longoni, C., Bonezzi, A. and Morewedge, C.K. (2019), "Resistance to medical artificial intelligence", *Journal of Consumer Research*, Vol. 46 No. 4, pp. 629-650.

Luo, X., Tong, S., Fang, Z. and Qu, Z. (2019), "Machines vs. humans: the impact of artificial intelligence chatbot disclosure on customer purchases", *Marketing Science*, Vol. 38 No. 6, pp. 937-947.

Mende, M., Scott, M.L., Van Doorn, J., Grewal, D. and Shanks, I. (2019), "Service robots rising: how humanoid robots influence service experiences and elicit compensatory consumer responses", *Journal of Marketing Research*, Vol. 56 No. 4, pp. 535-556.

Moon, J.W. and Kim, Y.G. (2001), "Extending the TAM for a World-Wide-Web context", *Information and Management*, Vol. 38 No. 4, pp. 217-230.

Morwitz, V.G., Steckel, J.H. and Gupta, A. (2007), "When do purchase intentions predict sales?", *International Journal of Forecasting*, Vol. 23 No. 3, pp. 347-364.

Mozafari, N., Weiger, W.H. and Hammerschmidt, M. (2020), "Resolving the chatbot disclosure dilemma: leveraging selective self-presentation to mitigate the negative effect of chatbot disclosure", *Proceedings of the Hawaii International Conference on System Sciences*, Vol. 54, pp. 2916-2923.

Mozafari, N., Weiger, W.H. and Hammerschmidt, M. (2022), "Trust me, I'm a bot – repercussions of chatbot disclosure in different service frontline settings", *Journal of Service Management*, Vol. 33 No. 2, pp. 221-245.

Munnukka, J., Talvitie-Lamberg, K. and Maity, D. (2022), "Anthropomorphism and social presence in human–virtual service assistant interactions: the role of dialog length and attitudes", *Computers in Human Behavior*, Vol. 135, p. 107343.

Murgia, A., Janssens, D., Demeyer, S. and Vasilescu, B. (2016), "Among the machines: human-bot interaction on social Q&A websites", *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Vol. 16, pp. 1272-1279.

Murtarelli, G., Gregory, A. and Romenti, S. (2021), "A conversation-based perspective for shaping ethical human–machine interactions: the particular challenge of chatbots", *Journal of Business Research*, Vol. 129 No. 5, pp. 927-935.

Pavlou, P.A. (2003), "Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model", *International Journal of Electronic Commerce*, Vol. 7 No. 3, pp. 101-134.

Pengnate, S.F. and Sarathy, R. (2017), "An experimental investigation of the influence of website emotional design features on trust in unfamiliar online vendors", *Computers in Human Behavior*, Vol. 67, pp. 49-60.

Peres, R., Schreier, M., Schweidel, D. and Sorescu, A. (2023), "On ChatGPT and beyond: how generative artificial intelligence may affect research, teaching, and practice", *International Journal of Research in Marketing*, Vol. 40 No. 2, pp. 269-275.

Pizzi, G., Scarpi, D. and Pantano, E. (2021), "Artificial intelligence and the new forms of interaction: who has the control when interacting with a chatbot?", *Journal of Business Research*, Vol. 129, pp. 878-890.

Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P. and Mazurek, G. (2019), "In bot we trust: a new methodology of chatbot performance measures", *Business Horizons*, Vol. 62 No. 6, pp. 785-797.

Qureshi, I., Fang, Y., Ramsey, E., McCole, P., Ibbotson, P. and Compeau, D. (2009), "Understanding online customer repurchasing intention and the mediating role of trust–an empirical investigation in two developed countries", *European Journal of Information Systems*, Vol. 18 No. 3, pp. 205-222.

Retail Week (2023), "Technology strategy at Zalando", available at: www.retail-week.com/prospect/zalando/technology-strategy-at-zalando/7038084.supplierarticle (accessed 24 May 2023).

Roskos-Ewoldsen, D.R. and Fazio, R.H. (1992), "On the orienting value of attitudes: attitude accessibility as a determinant of an object's attraction of visual attention", *Journal of Personality and Social Psychology*, Vol. 63 No. 2, pp. 198-211.

Roy, R. and Naidoo, V. (2021), "Enhancing chatbot effectiveness: the role of anthropomorphic conversational styles and time orientation", *Journal of Business Research*, Vol. 126, pp. 23-34.

Ryder, R.D. (2006), "Speciesism in the laboratory", in Singer, P. (Ed.), *In Defense of Animals: The Second Wave*, Blackwell Publishing, Malden, MA, pp. 87-103.

Sangle-Ferriere, M. and Voyer, B.G. (2019), "Friend or foe? Chat as a double-edged sword to assist customers", *Journal of Service Theory and Practice*, Vol. 29 No. 4, pp. 438-461.

Schellekens, G.A., Verlegh, P.W. and Smidts, A. (2010), "Language abstraction in word of mouth", *Journal of Consumer Research*, Vol. 37 No. 2, pp. 207-223.

Schmitt, B. (2020), "Speciesism: an obstacle to AI and robot adoption", *Marketing Letters*, Vol. 31 No. 1, pp. 3-6.

Schuetzler, R.M., Giboney, J.S., Grimes, G.M. and Nunamaker, J.F. (2018), "The influence of conversational agent embodiment and conversational relevance on socially desirable responding", *Decision Support Systems*, Vol. 114, pp. 94-102.

Sharma, S., Durand, R.M. and Gur-Arie, O. (1981), "Identification and analysis of moderator variables", *Journal of Marketing Research*, Vol. 18 No. 3, pp. 291-300.

Shawar, B.A. and Atwell, E.S. (2005), "Using corpora in machine-learning chatbot systems", *International Journal of Corpus Linguistics*, Vol. 10 No. 4, pp. 489-516.

Sheeran, P. (2002), "Intention—behavior relations: a conceptual and empirical review", *European Review of Social Psychology*, Vol. 12 No. 1, pp. 1-36.

Short, J., Williams, E. and Christie, B. (1976), *The Social Psychology of Telecommunications*, John Wiley and Sons, London.

Silva, S.C., De Cicco, R., Vlačić, B. and Elmashhara, M.G. (2023), "Using chatbots in e-retailing–how to mitigate perceived risk and enhance the flow experience", *International Journal of Retail and Distribution Management*, Vol. 51 No. 3, pp. 285-305.

Singer, P. and Mason, J. (2007), *The Ethics of What we Eat: Why Our Food Choices Matter*, Rodale Books, Emmaus, PA.

Skjuve, M., Følstad, A., Fostervold, K.I. and Brandtzaeg, P.B. (2021), "My chatbot companion - a study of human-Chatbot relationships", *International Journal of Human-Computer Studies*, Vol. 149, p. 102601.

Soares, A.M., Camacho, C. and Elmashhara, M.G. (2022), "Understanding the impact of chatbots on purchase intention", *Proceedings of the World Conference on Information Systems and Technologies*, Vol. 3, pp. 462-472.

Steinhoff, L., Arli, D., Weaven, S. and Kozlenkova, I.V. (2019), "Online relationship marketing", *Journal of the Academy of Marketing Science*, Vol. 47 No. 3, pp. 369-393.

Suh, B. and Han, I. (2003), "The impact of customer trust and perception of security control on the acceptance of electronic commerce", *International Journal of Electronic Commerce*, Vol. 7 No. 3, pp. 135-161.

Taylor, C.R., Kitchen, P.J., Sarkees, M.E. and Lolk, C.O. (2020), "Addressing the Janus face of customer service: a typology of new age service failures", *European Journal of Marketing*, Vol. 54 No. 10, pp. 2295-2316.

Thomaz, F., Salge, C., Karahanna, E. and Hulland, J. (2020), "Learning from the dark web: leveraging conversational agents in the era of hyper-privacy to enhance marketing", *Journal of the Academy of Marketing Science*, Vol. 48 No. 1, pp. 43-63.

Uysal, E., Alavi, S. and Bezençon, V. (2022), "Trojan horse or useful helper? A relationship perspective on artificial intelligence assistants with humanlike features", *Journal of the Academy of Marketing Science*, Vol. 50 No. 6, pp. 1153-1175.

Van der Goot, M.J., Koubayová, N. and Van Reijmersdal, E.A. (2024), "Understanding users' responses to disclosed vs. undisclosed customer service chatbots: a mixed methods study", *AI and Society*, Vol. 39 No. 6, pp. 2947-2960.

Van Doorn, J., Mende, M., Noble, S.M., Hulland, J., Ostrom, A.L., Grewal, D. and Petersen, J.A. (2017), "Domo arigato Mr. Roboto: emergence of automated social presence in organizational frontlines and customers' service experiences", *Journal of Service Research*, Vol. 20 No. 1, pp. 43-58.

Van Reijmersdal, E.A., Boerman, S.C. and van Noort, G. (2022), "Effects of online behaviorally targeted native advertising on persuasion: a test of two competing mechanisms", *Computers in Human Behavior Reports*, Vol. 7, p. 100221.

Wong, A. and Haque, M. (2022), "Understanding the brand and website effects of online loyalty: a mediation perspective", *Journal of Marketing Management*, Vol. 38 Nos 3/4, pp. 333-368.

Corresponding author
Maik Hammerschmidt can be contacted at: maik.hammerschmidt@wiwi.uni-goettingen.de