International Journal of Research in Marketing 41 (2024) 138–155

Contents lists available at ScienceDirect

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Sharing information with AI (versus a human) impairs brand trust: The role of audience size inferences and sense of exploitation

Deniz Lefkeli [a,*], Mustafa Karataş [b], Zeynep Gürhan-Canli [c]

[a] Department of Business and Management, LUISS Guido Carli University, Rome, Italy
[b] Graduate School of Business, Nazarbayev University, Astana, Kazakhstan
[c] Koç University, Istanbul, Turkiye

ARTICLE INFO

Article history:
First received on 19 July 2022 and was under review for 6½ months
Available online 1 September 2023

Keywords:
Consumer-technology interaction
Artificial intelligence
Information disclosure
Brand trust
Audience size
Sense of exploitation

ABSTRACT

This research examines whether and why disclosing information to AI as opposed to humans influences an important brand-related outcome—consumers' trust in brands. Results from two pilot studies and nine controlled experiments (n = 2,887) show that consumers trust brands less when they disclose information to AI as opposed to humans. The effect is driven by consumers' inference that AI shares information with a larger audience, which increases consumers' sense of exploitation. This, in turn, decreases their trust in brands. In line with our theorizing, the effect is stronger among consumers who are relatively more concerned about the privacy of their data. Furthermore, the negative consequences for brands can be mitigated when (1) customers are informed that the confidentiality of their information is protected, (2) AI is anthropomorphized, and (3) the disclosed information is relatively less relevant.
© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

"*Thank you for calling us. Please tell us your first pet's name so that we can verify your account.*" In the marketplace, consumers frequently have similar conversations with a human or an artificial intelligence (AI)-powered customer representative during which they are asked to disclose personal information such as birth date, contact details, financial information, social security number, or even the name of their first pet. In the last decade, companies have increasingly and rapidly transformed their operations, integrating AI into service for augmenting or substituting human employees (Hollebeek et al., 2021; Huang & Rust, 2018; McLeay et al., 2021; Ostrom et al., 2015; Van Doorn et al., 2017). Despite providing benefits, the new roles that AI assumes can incur costs not only for consumers but also for companies (Puntoni et al., 2021).

One of the many roles that AI performs in the marketplace is collecting data on customers. Although customers are often reluctant to share their personal information with AI or another human (Mothersbaugh et al., 2012), they are typically required to disclose information about themselves during their interactions with companies as a customer. How does this growing practice in the marketplace influence consumers' reactions to brands? More specifically, does disclosing information to AI have positive or negative outcomes for brands compared to disclosing information to humans? Although recent research exploring consumer-technology interaction has documented several antecedents of AI adoption (e.g., Longoni

* Corresponding author at: LUISS Guido Carli University, Viale Romania, 32, 00197 Rome, Italy.
E-mail addresses: dlefkeli@luiss.it (D. Lefkeli), mustafa.karatas@nu.edu.kz (M. Karataş), zcanli@ku.edu.tr (Z. Gürhan-Canli).

https://doi.org/10.1016/j.ijresmar.2023.08.011
0167-8116/© 2023 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

et al., 2019; Longoni & Cian, 2022) and factors that enhance consumers' information disclosure to companies (e.g., Lucas et al., 2014; Pickard et al., 2016; Uchida et al., 2017), relatively little is known about its consequences for brands. As one of the initial attempts to provide insights into the data collection role of AI in influencing company-level outcomes, we examine how consumers' trust in brands changes as a function of the agent to which they disclose information. By so doing, we identify a novel process based on a theorizing built on mind perception theory (Gray et al., 2007), which allows us to unveil managerial insights into the use of AI in information disclosure contexts.

Across two pilot studies and nine experiments, we show that customers' trust in brands decreases when they disclose information to AI. This effect is driven by customers' inference that AI would share the information they disclose with a larger audience, which, in turn, heightens their sense of exploitation. This heightened sense of exploitation, resulting from interactions with AI, impairs consumers' trust in brands. The negative consequences of disclosing information to AI are more pronounced among consumers who are highly concerned about the privacy of their data. The effect is mitigated when (1) contextual cues signal a smaller audience size (e.g., when confidentiality is assured), (2) AI is incorporated human-like features (i.e., when it is anthropomorphized), and (3) when the disclosed information is relatively less relevant.

Theoretically, our findings advance extant knowledge on customer-technology interactions in two significant ways. First, rather than examining how the type of agent (i.e., AI vs. human) impacts the propensity to disclose information (Kim et al., 2022; Lucas et al., 2014; Pickard et al., 2020; Pitardi et al., 2022; Schuetzler et al., 2018; Uchida et al., 2017), we explore the consequences of disclosing information to different agents. More specifically, we investigate how disclosing information to different agents influences consumers' trust in brands. In doing so, we shed light on the downstream consequences of disclosing information to AI for companies. Contributing to research on the impact of consumer-technology interaction on brand-related outcomes (Bergner et al., 2023; Cheng & Jiang, 2020; Cheng & Jiang, 2022; Lin and Wu, 2023; McLean et al., 2021; Minton et al., 2022), we identify brand trust as a novel consequence of disclosing information to different agents (i.e., AI vs. human). Second, we identify a novel psychological mechanism underlying the effect. As detailed in our theorizing, we base our conceptualization on the established finding in mind perception research (Gray et al., 2007) that consumers do not perceive AI as a social agent (e.g., Lefkeli et al., 2021). Thus, they anticipate a large audience as the recipients of their information when they disclose information to AI. This results in a higher sense of exploitation, which impairs brand trust. To the best of our knowledge, our research is the first to document that AI, as an interaction partner, changes consumers' audience size inferences. Further, we demonstrate how these inferences can influence individuals' sense of exploitation, which has not been explored in the context of consumer- technology interactions despite its pivotal importance for both companies and consumers (Puntoni et al., 2021). Substantively, by documenting theory-driven boundary conditions, we provide insights about how marketers could avoid the negative consequences of customers' interactions with AI in information disclosure contexts.

2. Theoretical background

AI appears in nearly all areas of consumers' lives, such as online shopping, banking, and the travel industry, and it has started replacing human employees in many industries. Although the use of AI is prevalent, consumers are still averse to using AI for various tasks that can be conventionally performed by humans (Castelo et al., 2019; Dietvorst et al. 2015; Longoni et al., 2019; Niszczota & Kaszas, 2020), a phenomenon known as algorithm aversion. Despite a general reluctance to prefer algorithms, recent research has identified several factors that would mitigate or attenuate algorithm aversion. Consumers are less reluctant to utilize AI, for example, when the decision task is objective (Castelo et al., 2019), when it is perceived as requiring consideration of unique characteristics of customers (Longoni et al., 2019), or when external factors make people think that human judgment is imperfect (Karataş & Cutright, 2023).

Although these findings significantly advance our understanding of consumer-technology interaction, they provide limited insight into the brand-related outcomes associated with disclosing information to AI or human agents. As shown in Table 1, prior research on information disclosure in contexts where consumers interact with technology has primarily focused on identifying antecedents of consumers' information disclosure and has explored how interacting with AI versus a human agent influences the propensity to disclose information (Kim et al., 2022; Lucas et al., 2014; Pickard et al., 2020; Pitardi et al., 2021; Uchida et al., 2017). On the other hand, emerging research that examined brand-related outcomes of consumers' interaction with technology (Bergner et al., 2023; Cheng & Jiang, 2020; Cheng & Jiang, 2022; Lin & Wu, 2023; McLean et al., 2021; Minton et al., 2022) has not examined information disclosure contexts. Additionally, these studies have largely focused on specific features and attributes of AI, without necessarily comparing consumer-human and consumer-technology interactions. In the current research, we address these gaps in the extant literature by examining the differential impact of disclosing information to an AI and a human agent on an important brand-related outcome—brand trust.

How might disclosing information to AI influence brand trust compared to disclosing information to humans? Drawing from established findings in cognitive psychology, which suggest that one's behaviors can influence their attitudes (e.g., Festinger, 1962), one might expect that disclosing information to AI would enhance consumers' trust in the brand. As individuals may be motivated to reduce the dissonance associated with sharing information with an agent they are averse to, it could lead to increased trust in the brand. On the other hand, research looking at the consequences of sharing information with AI and humans has found that emotional (i.e., feeling better, negative mood), relational (i.e., perceived warmth of the partner, interaction enjoyment, liking), and psychological (i.e., self-affirmation) outcomes of information disclosure to AI and

Table 1
Recent research on sharing information with AI.

Authors	Setting	Consumer Disclosure	Brand-related Outcomes	Comparing AI to Human(s)	Major Finding
Lucas et al., 2014	Information disclosure to virtual humans operated by computes vs actual humans	✓		✓	Interacting with a computer program as opposed to a human decreases fear of disclosure and impression management concerns as well as increasing the display of emotions and willingness to disclose.
Pickard et al., 2016	Preference for interacting with AI vs human in disclosure of sensitive topics	✓		✓	Individuals prefer human-like avatars controlled by computers over actual humans in disclosure of sensitive information because of the evaluative capability of the interviewer.
Uchida et al., 2017	Counseling system using human vs robots	✓		✓	Robots can promote disclosure of negative topics more than humans.
Ho et al., 2018	Consequences of self-disclosure to humans and robots	✓		✓	Emotional, relational, and psychological outcomes of disclosure are similar for humans and chatbots.
Schuetzler et al., 2018	Conversational agents' characteristics on disclosure	✓		✓	Individuals give more socially desirable responses to sensitive questions when they interact with humans compared to chatbots.
Pickard & Roster, 2020	The influence of the presence of human face on disclosure	✓		✓	Audio-only computer-assisted self-interview fosters more disclosure compared to a human-like agent or a human interviewer.
Cheng & Jiang, 2020	The impact of gratifications from chatbot use and privacy risks on customer satisfaction and brand loyalty		✓		Utilitarian (information), hedonic (entertainment), technology, and social gratifications from chatbot use increase satisfaction, while perceived privacy risk reduces it. Higher (vs. lower) levels of satisfaction in turn positively (vs. negatively) influences brand loyalty.
McLean et al., 2021	The attributes of AI voice assistants in influencing brand engagement		✓		The social presence, perceived intelligence, social attraction, usefulness, and ease of use of AI voice assistants positively influence consumer brand engagement. However, trust concerns have negative impact on brand engagement.
Pitardi et al., 2021	The effect of interaction partner on anticipated embarrassment	✓		✓	Interacting with robots as opposed to humans decreases anticipated embarrassment. The effect is driven by the lower levels of agency attribution to robots.
Srinivasan & Sarial-Abi, 2021	Role of algorithm vs human in a brand harm crisis		✓	✓	Consumers respond less negatively when a brand crisis is caused by AI as they attribute less agency and responsibility to AI compared to humans.
Cheng & Jiang, 2022	The impact of chatbot marketing efforts on customer response		✓		Interaction, information, accessibility, entertainment, and customization of chatbot service agents positively influence customer-brand relationships, which in turn enhance brand loyalty, preference for the brand, and purchase intention.
Kim et al., 2022	Consumer self-disclosure of sensitive information to AI/human	✓		✓	Consumers disclose sensitive information to AI more than they do to humans because of their lay beliefs about AI's incapability of social judgment.
Minton et al., 2022	Consumer religiosity and responses to AI and companies using AI		✓	✓	Consumer religiosity positively influences attitudes toward AI employees as well as companies using AI.
Lin & Wu, 2023	The role of perceived contingency of chatbots in driving consumer engagement and consumer responses		✓		Perceived contingency of social media chatbots enhances customer engagement, which then positively influences customer responses related to both the brand (i.e., intimacy, attachment, and purchase intention) and the chatbot (i.e., intention to continued use).
Bergner et al., 2023	The impact of AI conversational interfaces on consumer-brand relationship	✓	✓		Integrating key properties of human-to-human dialogue into interactions of consumers with AI promotes more intimate consumer-brand relationships and results in higher willingness to pay for the brand, higher brand advocacy and higher brand loyalty.
Current Research	The effect of information disclosure to AI vs human on brand related outcomes	✓	✓	✓	Disclosing information to AI as opposed to humans decreases consumers' trust in brands. The effect is driven by consumers' inference that an artificial agent shares information with a larger audience, which increases consumers' sense of exploitation.

humans do not differ (Ho et al., 2018). This finding then suggests that consumers may not differ in trusting a brand after disclosing information to AI or a human. In the current research, we build on the mind perception theory to hypothesize that sharing information with AI indeed impairs brand trust. In what follows, we briefly review the literature on the link between consumer information disclosure and brand trust, followed by how consumers perceive AI.

2.1. Consumer information disclosure and trust

Although trust has been defined and conceptualized from various perspectives in consumer research, these conceptualizations all incorporate consumers' confidence in the dependability (i.e., whether the brand has good intentions) and competence (i.e., whether the brand is capable of actualizing its intentions) of a brand as fundamental aspects of brand trust (Herbst et al., 2012). In this vein, these dimensions of trust could be mapped onto the fundamental dimensions of social perception in general, which have been named differently in various research streams such as warmth and competence (Fiske et al., 2007), nurturance and dominance (Wiggins & Broughton, 1991), communion and agency (Abele & Wojciszke, 2007), and agency and experience (Gray et al., 2007). As a crucial factor in forming sustainable and successful marketing relationships (Morgan & Hunt, 1994; Urban et al., 2000), trust has been considered an asset that is gradually built on former interactions with brands (Albert & Merunka, 2013).

Most pertinent to the current research, studies on brand trust have investigated the link between brand trust and disclosing information. These studies have consistently found that trust significantly influences consumers' information sharing, such that consumers are more willing to disclose information when they trust in a brand or company more (Dinev & Hart, 2006; Hoffman, et al., 1999). Furthermore, initial levels of trust established before a customer interacts with a brand enhance the disclosure of personal information (Dinev & Hart, 2006; Hoffman et al., 1999; Metzger, 2017; Treiblmaier & Chong, 2013; Wottrich et al., 2017). Relatedly, contextual cues implying trustworthiness, such as personalized social cues of immediacy (Lee & LaRose, 2011) and privacy seals (Keith et al., 2015), enhance voluntary information disclosure by customers.

The conceptualization of trust in the past literature as being *gradually* established after several customer—brand interactions (Doney & Cannon, 1997; Garbarino & Johnson, 1999) limits the relevance of past studies to a plethora of current business practices. Today, many companies require consumers to share personal information even before they form a relationship with the brand. For instance, a customer needs to respond to at least fifty questions about themselves to open an account and start using the dating application *OkCupid*. Similarly, consumers share their previous experiences with yoga and reveal their preferences for background voice and music even before creating an account on the *Down Dog* yoga application. Likewise, they give a lot of information, including their temporal focus, interpersonal relationships, and goals, to open an account on the self-care application *Fabulous*. These and many other examples point out to the fact that consumers often need to disclose information to brands even before they have a chance to build trust in the brand.

Established research on interpersonal trust has shown that people rely on superficial cues, such as assumed characteristics of the interaction partners (DeSteno et al., 2012; Yip & Schweitzer, 2015), to make inferences regarding the likelihood of the interaction partner to exploit the relationship (De Cremer, 1999; Weiss et al., 2018). For instance, men with a greater facial-width ratio are considered less trustworthy because their partners use this facial cue to perceive them as being more likely to exploit their relationship (Ormiston et al., 2017; Stirrat & Perrett, 2010). Likewise, non-verbal cues that signal a higher likelihood of reciprocity (and, thus, a lower likelihood of exploitation) in an economic game can increase trust (DeSteno et al., 2012). These findings are in line with the suggestions that trust is built on the anticipated risk of exploitation (Clark & Waddell, 1985; Deutsch, 1958), and the perceived integrity of the interaction partner plays a critical role in establishing trust (Mayer et al., 1995).

Synthesizing key findings and addressing the gaps in the literature, we explore information disclosure as an antecedent of consumers' trust in brands. As consumers use contextual cues to infer the risk of exploitation during their interactions (e.g., DeSteno et al., 2012; Ormiston et al., 2017), we predict that the agent to which consumers disclose information will lead to different outcomes for brands. This is because humans and AI are perceived differently on fundamental dimensions of social perception, resulting in differential levels of a sense of exploitation, which we discuss next.

2.2. Consumers' perception of (and, disclosure to) AI

Past research on customer-technology interaction has evinced that consumers' attributions to AI are strikingly different from attributions made to humans (Shank & DeSanti, 2018). Mind perception, a psychological process in which people attribute mental states such as reasoning, consciousness, or emotions to the objects (Gray et al., 2007; Premack & Woodruff, 1978), is considered an essential element of social life (Waytz et al., 2010). Multiple research streams in social cognition have explored the attribution of higher or lower levels of mind (i.e., anthropomorphization and dehumanization) and have demonstrated that people attribute mind not only to other humans but also to other living beings and objects (e.g., Gray et al., 2007). The attribution of mind to other social and non-social agents is influenced by various factors that could be related to the perceiver or the perceived, such as the motivation for social connection (Epley et al., 2007), likeableness (Kozak et al., 2006), and perceived similarity (Haslam, 2006).

Importantly, research has shown that people attribute significantly lower levels of mind to AI than they do to other people (e.g., Gray et al., 2007; Srinivasan & Sarial-Abi, 2021). In other words, AI is considered less of a social agent (Lefkeli et al.,

2021), lacking fundamental human capabilities of agency and experience (Srinivasan & Sarial-Abi, 2021) or warmth and competence (Luo et al., 2019; Zhang, Chen, & Xu, 2022). These dimensions relate to the two fundamental aspects of trusting a counterpart in a social interaction, namely, dependability and competence (Herbst et al., 2012; Kulms & Koop, 2018). As AI has been conceptualized as a mere technical tool that has the capacity to make objective and neutral decisions (Green & Viljoen, 2020), consumers do not consider it as an interaction partner with its own intentions or experiences, traits typically attributed to other humans and social agents (Lefkeli et al., 2021).

Further empirical support for the finding that customers perceive AI as lacking human capabilities also comes from research on consumer information disclosure. Investigating the impact of the type of interaction agent (i.e., AI vs. human) on information disclosure, studies have found that consumers are more willing to disclose information to AI, mostly because of its lack of human-like capabilities such as social judgment capability (Kim et al., 2022), evaluative capability (Pickard et al., 2016), and agency (Pitardi et al., 2021). Not acknowledging AI as a social actor, customers who interact with AI do not experience as many impression management concerns (Lucas et al., 2014) or embarrassment (Pitardi et al., 2021) as they typically do during their social interactions.

As AI is not perceived as a social agent, we argue that a communication episode in which consumers disclose information to AI will result in inferences regarding the presence of other social agents (i.e., humans). In other words, consumers perceive AI, about which they have limited knowledge and are trying to make sense of (Bonezzi et al., 2022; Puntoni et al. 2021; Querci et al., 2022; Rai, 2020), as operating primarily at the surface level. They believe that the AI will share the disclosed information with other agents possessing agency and experience (i.e., humans) who operate in the background. Thus, we predict that consumers will be more likely to think that there is a larger audience with whom their information will be shared.

This prediction is consistent with previous findings that information disclosure requests and the very act of disclosing information prompt consumers to think about the potential recipients of the information they disclose (Luo, 2002). In today's world, most consumers are concerned that their information could be shared with third parties (Anton et al., 2010). The significance of these concerns regarding the recipients of the information is evident in past findings, which demonstrate that signaling the brand's non-malicious intent by assuring that the information will not be shared with a large audience (Keith et al., 2015; Rifon et al., 2005) positively impacts consumer disclosure.

2.2.1. Information disclosure to AI and audience size inferences

To the best of our knowledge, no study has yet examined the impact of information disclosure on audience size inferences. Thus, we tested this critical conjecture of our theorizing in two pilot studies. In the first study, students from a European university (n = 164; M_{age} = 22 years; 92 female) participated in a single-factor (agent: AI vs. human) study in exchange for course credit. We randomly assigned participants to one of the two agent conditions. In the AI (vs. human) condition, participants recalled a time when they called a call center and a chatbot (vs. a customer representative) asked them to disclose personal information. They wrote down what they thought at that time.

We analyzed the data using OpenAI's GPT-3 (i.e., ChatGPT), a large language model (LLM) chatbot, which has been shown to be an acceptable method for coding open-ended responses (Mellon et al., 2022). GPT-3 categorized responses based on the inferences participants made about the recipients of the information (see Appendix A1 for details).

Overall, 73% of the participants made inferences about the potential recipients of the information shared with the customer service agent, which was significantly different from the test proportion of .5 ($p < .001$). The percentage of participants making an inference in both the human (78%; $p < .001$) and AI (68%; $p = .001$) conditions was significantly higher than the test proportion. Importantly, the proportion of participants who made an inference about audience size did not differ across the two conditions ($p = .12$).

In a separate pilot, we tested the second critical aspect of our theorizing regarding inferences about the potential audience. More specifically, we tested 1) whether inferences about the potential recipients of the information were peculiar to information disclosure (i.e., if customers were particularly prone to make an inference about audience size when they disclosed information compared to when they did not), and 2) if information disclosure to AI resulted in the inference of a larger audience compared to information disclosure to a human.

We recruited four hundred US-based participants (194 females, 3 non-binary; M_{age} = 43.8 years) on Amazon Mechanical Turk in return for a small monetary payment. The study employed a 2 (information disclosure: absent vs. present) × 2 (agent: human vs. AI), between-subjects design. Participants were randomly assigned to one of the two agent conditions and were informed about a fictitious internet service provider, *Conternet*, which was purportedly presented as a new company with special offers as part of its launch campaign. They were asked to imagine that they called the customer service of *Conternet* to learn more about their special deals. Participants in the AI (vs. human) condition imagined that they contacted an AI (vs. human) customer representative. Participants in the information disclosure condition were additionally asked to imagine that they provided their contact details during the call. Finally, participants rated the size of the audience with whom the content of this call would be shared (1: "a few people;" 7: "several people") and rated the customer service agent's warmth and competence (1: "not at all;" 7: "very much").

A two-way ANOVA on perceived audience size resulted in significant main effects of agent (M_{AI} = 3.78, SD = 2.12; M_{human} = 3.24, SD = 1.83; $F(1, 395) = 7.54$, $p = .006$) and information disclosure ($M_{no\ disclosure}$ = 3.73, SD = 2.10; $M_{disclosure}$ = 3.29, SD = 1.87; $F(1, 395) = 4.95$, $p = .027$), and an insignificant interaction term ($F(1, 395) = 1.63$, $p = .20$). Planned contrasts showed that perceived audience size was larger when people disclosed information to AI ($M_{AI-disclosure}$ = 3.68, SD = 2.08)

compared to when they disclosed to humans ($M_{human-disclosure}$ = 2.89, SD = 1.54; $F(1, 395) = 8.08$, $p = .005$). However, perceived audience size was not significantly different between AI and human conditions when information disclosure was absent ($M_{AI-no\ disclosure}$ = 3.87, SD = 2.17; $M_{human-no\ disclosure}$ = 3.58, SD = 2.03; $F(1, 395) = 1.08$, $p = .30$). Importantly, when customers imagined interacting with AI, the audience size perceptions did not differ between information disclosure ($M_{AI-disclosure}$ = 3.68, SD = 2.08) and no information disclosure conditions ($M_{AI-no\ disclosure}$ = 3.87, SD = 2.17, $F(1, 395) = .45$, $p = .5$). Supporting our theorizing, we found significant main effects of agent on competence (M_{AI} = 4.61, SD = 1.57; M_{human} = 5.23, SD = 1.18; $F(1, 395) = 19.22$, $p < .001$) and warmth (M_{AI} = 3.29, SD = 1.64, M_{human} = 4.96, SD = 1.14; $F(1, 395) = 138.18$, $p < .001$) perceptions. We found that AI was perceived as significantly less competent and less warm than the human representative, while the main effects of disclosure (p's > .38) and the interaction terms (p's > .53) were insignificant for both competence and warmth perceptions.

The significant main effect of information disclosure on audience size suggests that customers think about the potential parties that may have access to the information regarding their behavior to a greater extent, especially when they disclose information. Furthermore, they infer a larger audience in their interactions with AI, to which they ascribe lower levels of social capabilities of warmth and competence, compared to humans.

2.3. The current research

Building on the aforementioned literature and the findings of the two pilot studies, we argue that disclosing information to AI, as opposed to humans, will influence brand trust through differential levels of the sense of exploitation that will be experienced by customers. More specifically, we predict that disclosing information to AI will induce a higher sense of exploitation, driven by the inference that the information will be shared with a larger audience compared to disclosing information to humans. We further predict that this higher level of the sense of exploitation will decrease consumers' trust in brands.

Note that our main prediction is based on a process that involves the mediating role of two important factors, which points to theory-driven boundary conditions for the proposed effect. First, we predict and show in one pilot study that the effect is driven by an inference regarding the number of people with whom the disclosed information is shared by different agents (AI versus human). Specifically, the anticipated audience is larger in the case of sharing information with AI. This suggests that contextual cues signaling a smaller audience will mitigate the negative consequences of disclosing information to AI. Therefore, we further predict that an important and widely used contextual cue that signals a smaller audience size—informing participants that the information will not be shared with third parties—will attenuate the effect due to its impact on inferences regarding the anticipated audience size.

Likewise, certain elements such as the relevance of the requested information in a communication episode can increase the anticipated audience size even when consumers disclose information to humans. The level of mind attribution to a machine will not depend on the relevance of the information requested by AI. However, if a social actor (i.e., another human) requests the disclosure of irrelevant information, consumers will think that he or she is ill-intended and might share the information with others (see Martin et al., 2017 for similar propositions). Thus, we also predict that a request for irrelevant information made by humans will lead to lower trust in the brand.

The second mediator in our conceptualization is the sense of exploitation that will result from the anticipated audience size. Although past research documents that consumers are concerned about data security, confidentiality, and the use of their data by other parties (Lambrecht & Tucker, 2013; Morey et al., 2015; Steenkamp & Geyskens, 2006), some consumers are more susceptible to these negative consequences than others, especially when they share sensitive information (Milne et al., 2017). We predict that the sense of exploitation will impact brand trust even more intensely when consumers have higher levels of privacy concerns. However, this effect will be relatively weaker among those with lower levels of concern regarding the confidentiality and use of their information by third parties.

Finally, at the core of our theorizing lies the idea that consumers do not perceive AI as a social counterpart; instead, they attribute lower levels of mind and perceive AI as lacking human capabilities. This suggests that anthropomorphization of the agent can attenuate the effect, as it leads to attributing mind to AI (Kiesler et al., 2008) and results in perceiving AI as a social actor with human-like capabilities (Waytz et al., 2010; see Fig. 1 for a summary of our conceptual model).

2.4. Overview of studies

In the remainder of this article, we present results from seven studies that test our predictions. In Study 1A, using click-through rate as a proxy for trust, we demonstrate in a real-life setting that consumers trust brands less when they are informed that they will disclose information to an AI (vs. human) customer representative. In Study 1B (as well as Studies 1C and 1D, reported in Appendix A3), we use different types of information (i.e., financial, contact, and personally identifiable information) and brands (i.e., both fictitious and real) to show that consumers trust brands less after considering disclosing information to AI compared to humans. In Study 2, we provide direct empirical evidence for our proposed mechanism that disclosing information to AI increases the anticipated audience size, which, in turn, heightens the sense of exploitation and decreases consumers' trust in brands. In Study 3, we identify the moderating role of consumers' privacy concerns and show that the effect of disclosing information to AI on brand trust is more pronounced for individuals highly concerned about their privacy. Finally, in Studies 4–6, we identify theory-driven boundary conditions. Specifically, we show that the obtained

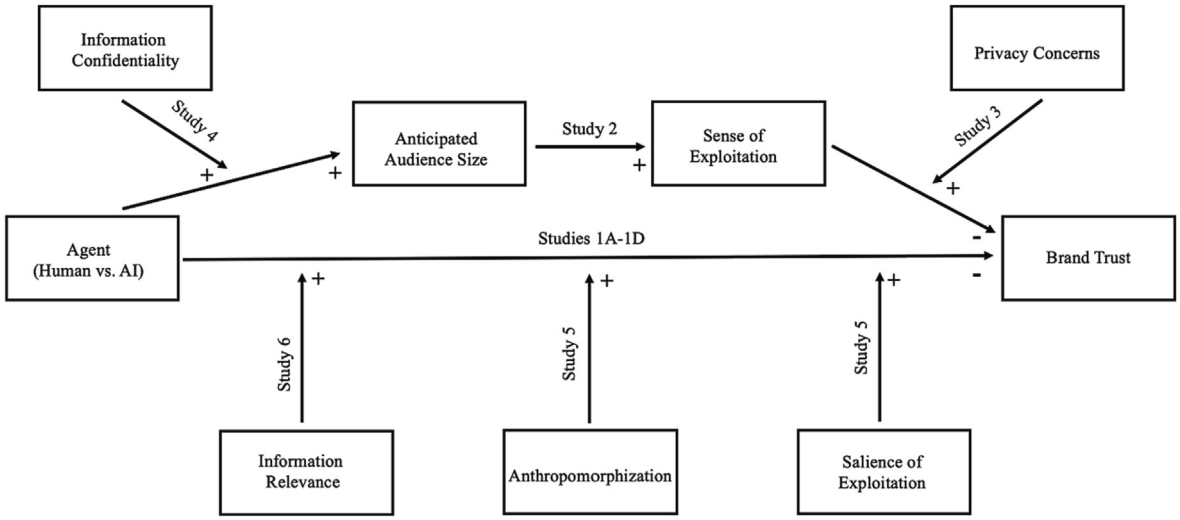


Fig. 1. The conceptual model.

effects are mitigated when (1) customers are informed that the confidentiality of their information is protected, (2) AI is anthropomorphized, and (3) the information is relatively less relevant.

3. Study 1A

Our main objective in Study 1A is to explore the impact of disclosing information to a human versus an AI agent on consumers' trust in brands in a real-world setting. With this goal, we conduct a field experiment on Google AdWords using advertisements that promote disclosing information to a human or an AI agent. We use clickthrough as an indicator of consumers' trust in brands (see Appendix A2 for a pretest). We predict that people will be less likely to click on a search ad that asks them to share personal information with an AI agent than one that asks sharing information with a human.

3.1. Method

We designed two versions of an online advertisement for a fictitious financial advisory firm, *Winvestus*. The headline in the AI condition read, "*Talk to our AI | Share financial information | Manage your money better*." This headline was accompanied by the information that read, "*Share your financial information with our artificial agent customer representatives. Our AI experts will use this information to best manage your money*." The headline in the human condition read, "*Talk to our people | Share financial information | Manage your money better*." Further, the information in the human condition read, "*Share your financial information with our human customer representatives. Our human experts will use this information to best manage your money*" (see Fig. 2 for sample ads used in the ad campaign).

Based on the recommendations of Google Ads, we specified seven keywords (e.g., financial advisor, investment management, money, profit) to advertise the search ads. In other words, one of the two ads was shown to Internet users in the United States whenever they made a Google search using these keywords. The ads were run over five days in June 2021 with an approximate daily campaign budget of 20 USD. Those who clicked on our ads were taken to a separate webpage explaining that the shown ad was part of an academic research.

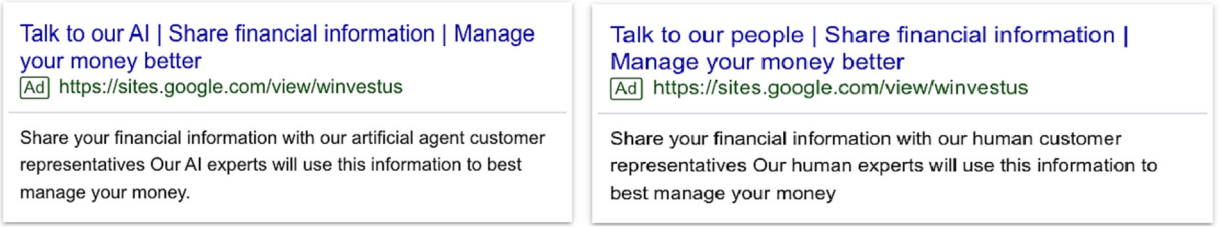


Fig. 2. Google Ad campaigns in Study 1.

3.2. Results and discussion

The two campaigns received a total of 2078 impressions and 31 clicks. In line with past research (Kronrod et al., 2012; Winterich et al., 2019), we used clickthrough rates (i.e., the average number of clicks per appearance) as our dependent measure. We conducted a chi-square analysis of the difference between the average percentage of the click-through on the AI and the human versions of the campaign. The analysis revealed that the click-through rate was significantly higher for the advertisement that promoted information disclosure to a human (1.97%) than for the advertisement that promoted information disclosure to AI (.88%; $\chi^2 = 4.14$, $p < .05$).

Involving a behavioral measure, this finding suggests that consumers trust a brand less when its ad implies disclosing information to AI as opposed to humans. Although it is a small-scale campaign, the significant difference between consumers' clicks on these ads provides support for our hypothesized effect. Furthermore, it is a stronger test for our prediction because past research has shown that consumers are not reluctant to use algorithms in objective tasks such as taking financial advice (Castelo et al., 2019). Our finding in Study 1A, which tested the hypothesized effect in a relatively objective, cognition-driven context, suggests that algorithm aversion can extend even to objective tasks when it requires consumers to disclose information.

4. Study 1B

In Study 1B, our aim is to test the predicted effect by directly measuring brand trust. We expect that consumers would report lower levels of brand trust after disclosing information to AI compared to a human representative.

4.1. Method

We recruited one hundred and ninety US-based participants (114 female; M_{age} = 45.3 years) on Amazon Mechanical Turk in exchange for a small monetary payment. Two participants who failed an initial attention check question were excluded from the final analysis. The study employed a single-factor (agent: AI vs. human), between-subjects design. Participants in the AI (vs. human) condition were asked to imagine that they called their bank to get information about the interest rates, and that they shared information regarding their savings with an AI (vs. human) customer representative. Then, participants completed the 4-item brand trust scale (Chaudhuri & Holbrook, 2001), which was extensively used in past research (Becerra & Badrinarayanan, 2013; Coyle et al., 2012; Herbst et al., 2012; Kim et al., 2019). Finally, participants indicated their demographics and were thanked.

4.2. Results and discussion

We first computed a composite brand trust score by averaging participants' responses to scale items (Cronbach's alpha = .96). A one-way ANOVA on this composite score revealed a significant effect of the agent ($F(1, 186) = 7.20$, $p < .01$). Supporting our basic prediction, participants who imagined disclosing financial information to an AI customer representative reported lower levels of trust in the brand (M_{AI} = 5.07, SD = 1.43; M_{human} = 5.60, SD = 1.26).

This finding provided further support for our prediction that disclosing information to AI results in lower levels of trust than disclosing information to a human. Moreover, two additional studies reported in Appendix A3 corroborated these findings by showing that the effect occurs also for fictitious brands with which customers did not have any prior interaction, and when they share different types of information (i.e., demographic and contact information).

5. Study 2

The main objective of Study 2 is to explore the underlying mechanism of the effect of agent type on brand trust. We propose that when consumers are asked to disclose their information, they will think that AI will share this information with a larger number of people compared to a human, which will heighten their sense of exploitation. We further predict that this heightened sense of exploitation will result inlower levels of trust in brands. Study 2 directly tests this serial mediation model. The second goal of Study 2 is to rule out an alternative explanation based on the perceived value of information. It might be argued that consumers infer that algorithms are programmed to obtain more valuable information for companies, which could alternatively explain the observed effect. The mind perception theory, on the other hand, does not imply that the value of the information to be shared with AI (i.e., whether relatively less or more valuable information is shared) would change the perception that AI lacks its own intentions. Thus, we expect our findings to hold irrespective of the perceived value of the information.

5.1. Method

We recruited three hundred and two US-based participants (154 female, 2 non-binary; M_{age} = 41.4 years) on Amazon Mechanical Turk in return for a small monetary payment. Participants were randomly assigned to one of two agent

conditions (agent: AI vs. human) and were informed about a new brand entering the market. We then asked participants to imagine that they shared their contact information with either a human or an AI customer representative. After reading the text, participants completed the brand trust scale (Chaudhuri & Holbrook, 2001). We next measured anticipated audience size (adapted from Barasch & Berger, 2014) by asking participants to rate the number of people that the agent would share the information they disclosed on a 7-point scale (1 = "a few people," 7 = "several people"), using the same single-item measure as in pilot study 2. Then, we asked participants two questions measuring the extent to which they believe that the representative would be exploitative and willing to take advantage of their information (1 = "not at all," 7 = "to a great extent;" Clark & Waddell, 1985). Next, we measured perceived value of the information by asking participants two questions on their perception of the importance and meaningfulness (1 = "not at all," 7 = "to a great extent;" adapted from Zhang et al., 2011) of the information they shared with the agent. Finally, participants provided demographic information and were thanked.

5.2. Results

We excluded the response of one participant who failed an initial attention check question, resulting in a final sample size of three hundred and one.

Brand Trust. A one-way ANOVA resulted in a significant main effect of the agent ($F(1, 299) = 9.93$, $p = .002$) on the composite brand trust score (Cronbach's alpha = .95). Participants who imagined disclosing contact information to AI trusted the brand less than those who imagined disclosing it to a human ($M_{AI} = 4.04$, SD = 1.40; $M_{human} = 4.54$, SD = 1.36).

Perceived Value of Information. A one-way ANOVA yielded that the agent with which participants imagined sharing information did not influence the perceived value of the information ($M_{AI} = 4.58$, SD = 1.69; $M_{human} = 4.56$, SD = 1.57; $F(1, 299) = .009$, $p > .9$). Importantly, we checked the potential mediating role of the perceived value of information using PROCESS (Hayes, 2018) Model 4 with 10,000 bootstrapped samples. The test yielded that the indirect effect of agent type on brand trust through the perceived value of information was insignificant ($\beta = -.002$, $CI_{95\%} = [-.06, .06]$). These findings rule out the possibility that the effect could be explained by differences in the perceived value of information.

Serial Mediation Analysis. We tested the proposed mediation model using PROCESS (Hayes, 2018) Model 6 with 10,000 bootstrapped samples. In the model, we included agent type (0: human, 1: AI) as the independent variable, brand trust as the dependent variable, anticipated audience size as the proximal mediator, and the sense of exploitation as the distal mediator. Serial mediation analysis showed that the agent type significantly increased the anticipated audience size ($\beta = .58$; $t(299) = 2.89$, $p = .004$; $CI_{95\%} = [.19, .98]$), which in turn significantly increased the consumers' sense of exploitation ($\beta = .72$; $t(298) = 20.12$, $p < .001$; $CI_{95\%} = [.65, .79]$). This sense of exploitation then significantly decreased consumers' trust in brands ($\beta = -.38$; $t(297) = -6.11$, $p < .001$; $CI_{95\%} = [-.51, -.26]$).

The hypothesized indirect effect of agent type on brand trust through anticipated audience size and the resulting sense of exploitation was significant ($\beta = -.16$, $CI_{95\%} = [-.29, -.05]$). The direct effect of the agent type on brand trust was reduced after accounting for the proposed mediators but remained significant ($\beta = -.35$; $t(297) = -2.52$, $p = .012$; $CI_{95\%} = [-.61, -.08]$), suggesting partial mediation (see Fig. 3 for path coefficients).

Importantly, we found that the models with either of the two proposed mediators were insignificant. More specifically, the effect of agent type on brand trust was not mediated by only the audience size inferences ($\beta = -.04$, $CI_{95\%} = [-.14, .02]$) or by only the sense of exploitation ($\beta = .05$, $CI_{95\%} = [-.05, .14]$). Moreover, a serial mediation model in which we reversed the order of the mediators (i.e., agent type → sense of exploitation → brand trust) was not significant ($\beta = -.02$, $CI_{95\%} = [-.07, .01]$).

5.3. Discussion

Study 2 replicated our main finding that disclosing information to AI results in lower trust in brands than disclosing information to a human. More importantly, results from Study 2 provided empirical evidence in support of our process prediction. We showed that the negative effect of sharing information with AI on brand trust is driven by an inference regarding the anticipated size of the audience with which this information will be shared, and subsequently, by the resulting sense of exploitation.

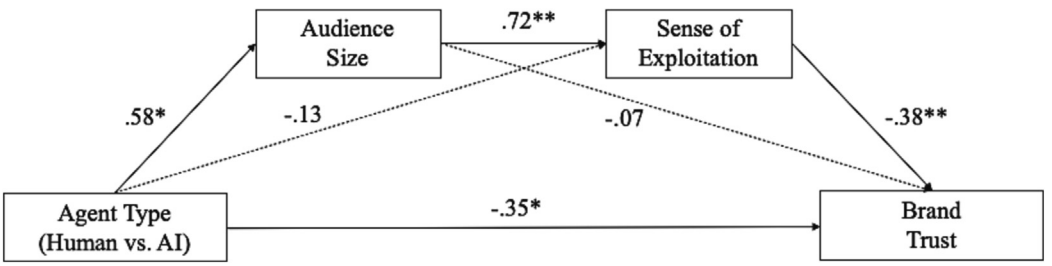


Fig. 3. Serial mediation model in study 2. Note.—Path coefficients of serial mediation model in Study 2. *p <.05; **p <.001.

6. Study 3

The main objective of Study 3 is to test the hypothesized role of individual differences in privacy concerns. If our theory holds true, then the sense of exploitation should have a greater impact on brand trust among consumers who are relatively more concerned about the privacy of their data. To test this prediction, we measure participants' chronic levels of privacy concerns in Study 3 and test their role in the hypothesized model. Another goal is to enhance the generalizability of our process evidence, which we documented in Study 2, by testing it in another context that involves sharing a different type of information.

6.1. Method

Two hundred US-based participants (116 female; M_{age} = 47.3 years), recruited on Amazon Mechanical Turk, completed the study in exchange for monetary payment. Participants were randomly assigned to one of the two conditions (agent: AI vs. human). Participants in the AI (vs. human) condition were asked to imagine that they had called their bank to learn about a new credit card, and that they had provided their personal details, including their date of birth, social security number, phone number, e-mail address, home address, and their current credit card limit, to an AI (vs. human) customer representative.

Next, akin to Study 2, participants completed the same brand trust scale (Chaudhuri & Holbrook, 2001), rated the size of the anticipated audience (Barasch & Berger, 2014), and reported their sense of exploitation (Clark & Waddell, 1985). Finally, we measured participants' concerns regarding organizational information privacy practices (Smith et al., 1996).

6.2. Results

We excluded one participant who failed an initial attention check question, resulting in a final sample of one hundred and ninety-nine participants.

Brand Trust. Replicating our findings from previous studies, a one-way ANOVA revealed a significant main effect of agent ($F(1, 197) = 5.94$, $p = .016$) on brand trust (Cronbach's alpha = .97). Participants who imagined disclosing information to AI trusted the brand less (M_{AI} = 4.81, SD = 1.69; M_{human} = 5.34, SD = 1.35).

Moderated Serial Mediation. We tested the proposed moderated serial mediation model using PROCESS (Hayes, 2018) Model 87 with 10,000 bootstrapped samples. We obtained a significant moderated mediation index (index = -.09, $CI_{95\%}$ = [-.19, -.01]). The analysis showed that disclosing information to AI increased the anticipated audience size (β = .90; $t(197) = 3.76$, $p < .001$; $CI_{95\%}$ = [.43, 1.37]), which then significantly increased the sense of exploitation (β = .70; $t(196) = 16.12$, $p < .001$; $CI_{95\%}$ = [.62, .79]).

This heightened sense of exploitation interacted with participants' chronic levels of privacy concerns to predict their trust in the brand. Importantly, the direct effect of the sense of exploitation on brand trust was no longer significant when the level of privacy concerns was included as a moderator (β = .39; $t(193) = 1.29$, $p = .19$; $CI_{95\%}$ = [-.21, .99]). Similarly, the direct effect of privacy concerns was only marginally significant (β = .34; $t(193) = 2.01$, $p > .05$; $CI_{95\%}$ = [-.005, .69]). The interaction between the sense of exploitation and privacy concerns, however, significantly influenced brand trust (β = -.14; $t(193) = -2.70$, $p < .01$; $CI_{95\%}$ = [-.24, -.04]). As predicted, the conditional effect was stronger among those with a higher sense of privacy concerns (β = -.55; $t(193) = -6.52$, $p < .001$; $CI_{95\%}$ = [-.72, -.38]) than those who were relatively less concerned about their privacy (β = -.31; $t(193) = -4.14$, $p < .001$; $CI_{95\%}$ = [-.46, -.16]; see Fig. 4 for path coefficients).

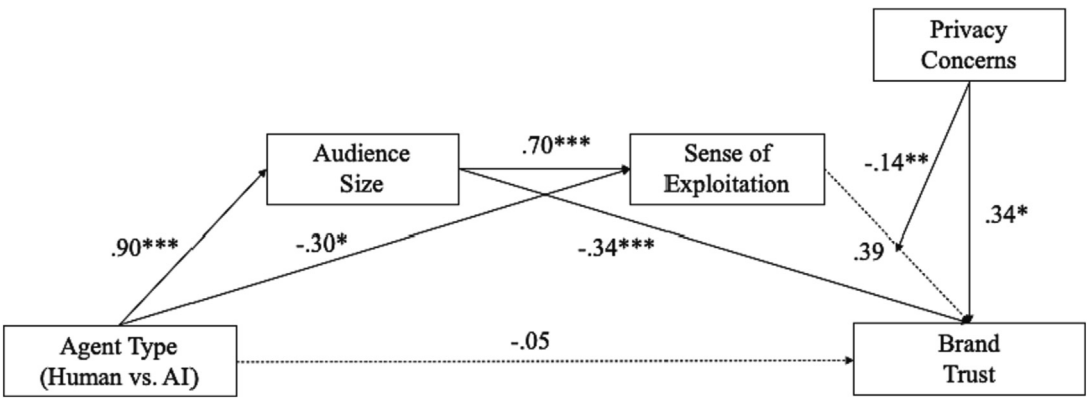


Fig. 4. Moderated mediation model in study 3. Note.—Path coefficients of moderated mediation model in Study 4. *p <.05; **p <.01; ***p <.001.

6.3. Discussion

Replicating previous findings, these results show that disclosing information to AI increases the anticipated audience size, evoking a higher sense of exploitation, which, in turn, decreases consumers' trust in brands. Study 3 also provides further support for our process prediction by demonstrating that brand trust decreases even more among consumers who are more sensitive to the privacy of their data when they feel exploited.

7. Study 4

The objective of Study 4 is to test another moderator based on our theorizing: informing customers about the confidentiality of their information. Note that our theorizing predicts that the effect is driven by an initial inference regarding the size of the audience with whom this information would be shared. According to this proposed process, the effect should mitigate when contextual cues signal a lower number of third parties with whom the information could potentially be shared. Thus, we expect the effect to mitigate among participants who are informed that the confidentiality of the information will be protected.

7.1. Method

Study 4 employed a 2 (agent: AI vs. human) × 2 (confidentiality of the information: control vs. confidential), between-subjects design. We recruited four hundred and ninety-five US-based participants (254 female; M_{age} = 43.5 years) on Amazon Mechanical Turk in exchange for a small payment.

Participants were randomly assigned to one of the four conditions and read a scenario about a fictitious financial advising company. The scenario asked them to imagine sharing details about their financial savings with either a human or an AI representative. In the control condition, no additional information was provided. In the confidentiality condition, they were asked to imagine that the call was end-to-end encrypted, and that the details of the call would not be shared with third parties.

Next, we measured participants' trust in the brand (Chaudhuri & Holbrook, 2001), sense of exploitation (Clark & Waddell, 1985), and the size of the anticipated audience (Barasch & Berger, 2014). Finally, we asked participants to indicate their demographics.

7.2. Results

We excluded responses from 11 participants who failed an initial attention check question, resulting in a final sample of four hundred and eighty-four participants.

Audience Size. A two-way ANOVA revealed a significant interaction between agent type and the confidentiality of the information on perceived audience size ($F(1, 480) = 3.99$, $p = .046$). Pairwise comparisons showed that participants in the control condition thought that the information they disclosed would be shared with fewer people when the customer service agent was human ($M_{AI-control} = 3.98$, SD = 1.73; $M_{human-control} = 3.13$, SD = 1.70; $F(1, 480) = 16.49$, $p < .001$). However, the difference between the inferred audience size was mitigated when participants were informed that the information they shared would be kept confidential ($M_{AI-confidential} = 3.11$, SD = 1.75; $M_{human-confidential} = 2.86$, SD = 1.45; $F(1, 480) = 1.41$, $p = .24$). Importantly, assuring confidentiality significantly decreased the anticipated audience size among participants who imagined interacting with an AI agent ($M_{AI-confidential} = 3.12$, SD = 1.75; $M_{AI-control} = 3.98$, SD = 1.73; $F(1, 480) = 15.76$, $p < .001$). The difference between the two human conditions was not significant ($M_{human-confidential} = 2.86$, SD = 1.45; $M_{human-control} = 3.13$, SD = 1.70; $F(1, 480) = 1.53$, $p = .22$).

Brand Trust. A two-way ANOVA revealed significant main effects of agent ($M_{AI} = 4.31$, SD = 1.42, $M_{human} = 4.97$, SD = 1.25; $F(1, 480) = 30.56$, $p < .001$) and confidentiality of the information ($M_{control} = 4.40$, SD = 1.37; $M_{confidential} = 4.87$, SD = 1.35; $F(1, 480) = 15.15$, $p < .001$) on the composite brand trust score (Cronbach's alpha = .96). More importantly, we obtained a marginally significant interaction ($F(1, 480) = 3.86$, $p = .05$). Simple contrast analyses showed that participants in the control condition reported significantly lower levels of trust in brands when they imagined disclosing information to AI ($M_{AI-control} = 3.96$, SD = 1.34; $M_{human-control} = 4.85$, SD = 1.26; $F(1, 480) = 27.61$, $p < .001$). Moreover, participants who imagined interacting with AI reported significantly higher levels of trust when they were informed that the confidentiality of the information would be protected ($M_{AI-control} = 3.96$, SD = 1.34; $M_{AI-confidential} = 4.66$, SD = 1.43; $F(1, 480) = 17.15$, $p < .001$). Additionally, the difference between the levels of trust reported by participants in different agent type conditions was mitigated when they were informed about the protection of confidentiality ($M_{AI-confidential} = 4.66$, SD = 1.43; $M_{human-confidential} = 5.09$, SD = 1.24; $F(1, 480) = 6.46$, $p = .011$).

Moderated Serial Mediation. We tested the proposed moderated serial mediation model using PROCESS (Hayes, 2018) Model 83 with 10,000 bootstrapped samples. In the model, we included agent type (0: human, 1: AI) as the independent variable, information confidentiality as the moderating variable (0: confidentiality, 1: control), brand trust as the dependent variable, anticipated audience size as the proximal mediator, and the sense of exploitation as the distal mediator. The analysis revealed that the hypothesized model was significant (index = -.13, $CI_{95\%}$= [-.27, -.001]).

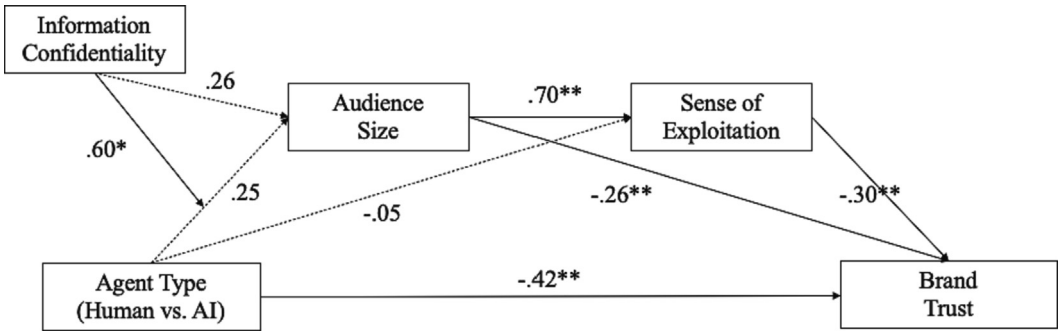


Fig. 5. Moderated serial mediation in study 4. note.—Path coefficients for the moderated serial mediation model in Study 4. *p <.05; **p <.001.

As we hypothesized, the indirect effect of the agent type on brand trust through the hypothesized moderator and mediators was significant only in the control condition (β = -.18; $CI_{95\%}$ = [-.30, -.08]), and it disappeared when participants were informed about the confidentiality of the information (β = -.05; $CI_{95\%}$ = [-.15, .03]). This was driven by the moderating role of confidentiality on the impact of agent type on audience size inferences. While participants in the control condition reported a significantly higher number of people with whom the information would be shared by AI than by a human (β = .86; t (480) = 3.97, p < .001; $CI_{95\%}$ = [.43, 1.28]), this difference disappeared in the confidentiality condition (β = .25; t (480) = .24, p = .24; $CI_{95\%}$ = [-.16, .67]). As in Studies 2 and 3, the anticipated audience size then significantly influenced the sense of exploitation (β = .70; t(480) = 24.95, p < .001; $CI_{95\%}$ = [.64, .75]), which in turn decreased trust in the brand (β = -.30; t(480) = -6.65, p < .001; $CI_{95\%}$ = [-.39, -.21], see Fig. 5).

7.3. Discussion

The findings of Study 4 provided additional support for our theorizing by showing that the lower sense of trust in brands after disclosing information to AI is mitigated when consumers are informed that the confidentiality of their information is protected. This is because protecting the confidentiality of the disclosed information significantly reduced the inferred audience size, leading to a lower sense of exploitation.

8. Study 5

The objectives of Study 5 are threefold. First, using a moderation-by-process approach (Spencer et al., 2005), we aim to directly test the mediating role of the sense of exploitation by manipulating it. By so doing, we provide more direct evidence for our proposed process.

Second, we aim to test the hypothesized moderation by anthropomorphism. As we predict and have shown in one preliminary pilot study, consumers do not consider AI to be a social agent compared to humans, which underlies the obtained effect. To directly assess the role of the extent to which AI is perceived as a social agent, we manipulate the anthropomorphism of AI. Given past research evincing that anthropomorphism results in attributing human-like qualities to AI (Gong, 2008; Eyssel & Kuchenbrandt, 2012), we predict the effect to attenuate when AI is anthropomorphized. Finally, to heighten ecological validity, we use a real behavioral measure in this study.

8.1. Method

Study 5 employed a 3 (agent: human vs. anthropomorphized AI vs. non-anthropomorphized AI) × 2 (salience of exploitation: low vs. high), between-subjects design. We recruited three hundred US-based participants (86 females, 10 non-binary; M_{age} = 33.3 years) on Prolific in return for monetary payment. We excluded three participants due to failure in an attention check question, resulting in a final sample size of two hundred and ninety-seven participants. Participants were randomly assigned to one of the six conditions. In the first part of the study, participants in the low salience condition were asked to recall a time when they shared information with a company and to briefly describe the experience. Participants in the high salience condition were asked to think about a time when they shared an information with a company and felt exploited. Next, they were given the description of a new dating app, *Matchify*, which was purportedly presented as being developed. In the human condition (vs. the two AI conditions), the application was described as involving a human relationship expert (vs. an algorithm), who would identify perfect matches for users based on their responses to 40 questions during account setup. We used different visuals to manipulate the anthropomorphization of AI (see Fig. 6). After reading this information about *Matchify*, participants reported their trust in the brand based on their impressions. Then, we administered our behavioral measure by asking participants if they were willing to receive a download link for the beta version of *Matchify* to try and



Fig. 6. Visuals used in the scenario for human, anthropomorphized AI, and AI.

provide feedback on the app (1 = "yes;" 2 = "no"). Finally, we asked a hypothesis guess question, but none of the participants were able to guess it.

8.2. Results

Brand Trust. A two-way ANOVA revealed significant main effects of agent (M_{AI} = 3.46, SD = 1.11; $M_{Anthropomorphized\ AI}$ = 4.19, SD = 1.26; M_{human} = 4.03, SD = 1.35; $F(2, 291) = 9.12$; $p < .001$) and salience of exploitation ($M_{low\ salience}$ = 4.08, SD = 1.32; $M_{high\ salience}$ = 3.68, SD = 1.19; $F(1, 291) = 6.77$, $p = .01$) on the composite brand trust score (Cronbach's alpha = .92). More importantly, it revealed a significant interaction of agent type and salience of exploitation on the composite brand trust score ($F(2, 291) = 6.95$, $p = .001$). Contrasts showed that when the sense of exploitation was not salient, participants in the non-anthropomorphized AI condition reported significantly lower levels of trust compared to participants in the anthropomorphized AI condition ($M_{non\text{-}anthropomorphized\ AI}$ = 3.27, SD = 1.10, $M_{anthropomorphized\ AI}$ = 4.48, SD = 1.27, $p < .001$) and participants in the human condition (M_{human} = 4.39, SD = 1.25, $p < .001$). The difference between the anthropomorphized AI and the human conditions was not significant ($p = .7$).

However, when the sense of exploitation was salient, the difference across groups was not significant ($F(2, 291) = .59$, $p = .56$). Participants in the human condition (M_{human} = 3.57, SD = 1.34) reported comparable levels of trust to participants in the non-anthropomorphized AI ($M_{non\text{-}anthropomorphized\ AI}$ = 3.64, SD = 1.10; $p = .79$) and the anthropomorphized AI ($M_{anthropomorphized\ AI}$ = 3.84, SD = 1.16; $p = .3$) conditions.

The effect was attenuated in the low salience condition because brand trust significantly decreased in the human ($M_{low\ salience—human}$ = 4.39, SD = 1.25, $M_{high\ salience—human}$ = 3.57, SD = 1.34; $p < .001$) and anthropomorphized AI ($M_{low\ salience—anthropomorphized\ AI}$ = 4.48, SD = 1.27, $M_{high\ salience—anthropomorphized\ AI}$ = 3.84, SD = 1.16; $p < .01$) conditions. The difference in brand trust across low and high salience conditions when the AI was non-anthropomorphized, however, did not reach significance ($M_{low\ salience—non\text{–}\ anthropomorphized\ AI}$ = 3.27, SD = 1.10, $M_{high\ salience—non\text{-}anthropomorphized\ AI}$ = 3.64, SD = 1.10, $p = .13$).

Behavioral Measure. When the AI was anthropomorphized, a significantly higher proportion of participants in the low salience condition was willing to get the beta version of the app than in the high salience condition ($\chi^2(1, 98) = 6.78$, $p < .01$). Similarly, in the human condition, more participants in the low salience condition were willing to get the beta version than in the high salience condition ($\chi^2(1, 97) = 6.33$, $p = .012$). We found no significant difference between the low and high salience conditions when the AI was not anthropomorphized ($\chi^2(1, 98) = .06$, $p > .8$).

8.3. Discussion

The findings of Study 5 provided compelling evidence for our proposed model. First, we presented direct evidence for the role of the sense of exploitation in our theorizing. By manipulating the sense of exploitation directly, we demonstrated that the effect was mitigated when the salience of exploitation was high. Second, in line with our conceptualization based on the mind perception theory, a human-like AI agent mitigated the negative impact of disclosing information on brand trust. Finally, by using a behavioral measure, we enhanced the ecological validity of our findings and showed that people were equally willing to use a product when they imagined disclosing information to a human or to an anthropomorphized AI. These preferences were higher than those observed in the non-anthropomorphized AI condition.

9. Study 6

Two objectives guide Study 6. First, we aim to provide further support for our theorizing as it relates to consumers' perception of AI as less of a social agent. More specifically, we test the moderating role of a variable which should not significantly impact the extent to which AI is attributed mind: the relevance of the requested information. We reason that AI will be perceived as lacking human capabilities, regardless of whether the information it requests is relevant or irrelevant. Consequently, we expect to obtain no significant difference in brand trust when AI requests relevant or irrelevant information. However, the relevance of the information is likely to influence the effect of disclosing information to humans as it will raise doubts about the use of this information (e.g., Martin et al., 2017), potentially resulting in differential levels of sense of

exploitation. Thus, in the human condition, we expect to obtain a significant difference based on the relevance of the information. Study 6 tests these predictions.

Second, unlike the studies we reported so far, Study 6 employs a relatively less subtle manipulation. Rather than informing participants about the characteristics of the information they provided, we directly ask questions that are either relevant or irrelevant to the decision context. This allows us to manipulate information disclosure by making participants provide information about themselves rather than merely imagining disclosing information.

9.1. Method

Study 6 employed a 2 (agent: AI vs. human) × 2 (information: relevant vs. irrelevant), between-subjects design. We recruited four hundred US-based participants (192 females, 3 non-binary; M_{age} = 37.9 years) on Prolific in return for monetary payment. We excluded three participants due to failure in an attention check question, resulting in a final sample size of three hundred and ninety-seven. Participants were randomly assigned to one of four conditions and were informed about a new app, *Vocalify*, which was purportedly presented as providing music recommendations based on customer input. In the AI (vs. human) condition, participants were informed that an algorithm (vs. expert musicians) would evaluate their responses and recommend a song. Next, participants in the relevant condition were asked to report their favorite music genre, their motivation for listening to music, the average duration they spend to listen to music on a typical day, and whether they could play a musical instrument. Participants in the irrelevant condition were asked to report their favorite music genre, their motivation for going on a vacation, the average duration they spend on the Internet on a typical day, and whether they had a life insurance. A pretest (n = 98; 58 females, 1 non-binary; M_{age} = 42.1) confirmed that the manipulation was successful. One-sample t-tests showed that relevance ratings for the relevant questions were significantly higher than the midpoint (M = 4.93, SD = 1.22; $t(97) = 7.52$, $p < .001$), while the ratings for the irrelevant questions were significantly lower than the midpoint (M = 2.03, SD = 1.47; $t(97) = -13.26$, $p < .001$).

Next, participants were asked to wait while either the algorithm or an expert musician was evaluating their responses. In the meantime, they were asked to complete the brand trust scale based on their impression of *Vocalify*. Finally, participants received a music recommendation and were asked to evaluate it by reporting their liking and their willingness to explore similar songs on a 7-point scale.

9.2. Results

Brand Trust. A two-way ANOVA revealed a significant main effect of relevance on the composite brand trust score (Cronbach's alpha = .82; $M_{relevant}$ = 4.14, SD = .78; $M_{irrelevant}$ = 3.96, SD = .87; $F(1, 393) = 4.35$, $p = .038$). The main effect of the agent (M_{AI} = 3.98, SD = 0.84; M_{human} = 4.11, SD = .80; $F(1, 393) = 2.90$, $p = .09$) and the interaction ($p < .08$) were marginally significant. Planned contrasts showed that participants reported significantly lower levels of trust when they disclosed relevant information to AI than when they disclosed relevant information to a human ($M_{relevant-AI}$ = 3.99, SD = .70; $M_{relevant-human}$ = 4.28, SD = .79; $F(1, 393) = 5.55$, $p < .02$). The difference between the levels of trust was mitigated when they were asked to report irrelevant information ($M_{irrelevant-AI}$ = 3.97, SD = .94; $M_{irrelevant-human}$ = 3.96, SD = .78; $F(1, 393) = .003$, $p = .97$). This effect occurred because trust within the human condition significantly dropped when the information was irrelevant ($F(1, 393) = 7.50$, $p = .006$). However, the difference between the levels of trust did not vary across AI conditions ($F(1, 393) = .048$, $p > .8$).

Recommendation Evaluation. The effect we obtained for trust did not spillover to actual evaluations. A two-way ANOVA revealed that the agent type ($F(1, 393) = 2.18$, $p = .14$), the relevance of the information ($F(1, 393) = 1.13$, $p = .29$), or their interaction ($F(1, 393) = 1.13$, $p = .14$) did not influence the liking of the recommended song. Similarly, the agent type ($F(1, 393) = .30$, $p = .59$), the relevance of information ($F(1, 393) = 1.74$, $p = .19$), or their interaction ($F(1, 393) = .39$, $p = .53$) did not significantly impact participants' willingness to explore similar songs.

9.3. Discussion

The findings of Study 6 showed that the relevance of information moderated the impact of agent type on brand trust. In line with our theorizing, when participants were asked to provide irrelevant information by humans, their trust in the brand significantly decreased.

10. General discussion

Across seven reported experiments (and, two studies reported in the Appendix A3), we showed that disclosing personal information to AI results in lower levels of brand trust than disclosing information to a human agent. This effect is driven by inferences that AI will share the information with a larger audience, which then heightens consumers' sense of exploitation. In line with our theorizing, we showed that this effect is mitigated when consumers are informed that their information will be kept confidential, when AI is anthropomorphized, and when the requested information is relatively irrelevant.

10.1. Theoretical contributions and future research

Our findings advance the growing stream of research on consumer-technology interactions in significant ways. Extant research has shown that consumers are averse to using algorithms for tasks that can be performed by humans (Dietvorst et al., 2015). These studies on algorithm aversion, however, mostly focused on technology adoption and examined when consumers prefer interacting with AI versus a human (e.g., Castelo et al., 2019; Longoni et al., 2019). We extend these studies by exploring the downstream consequences of forced interactions with different types of agents in an information disclosure context, which is increasingly becoming a widespread marketplace practice.

As discussed previously and summarized in Table 1, we add to studies on consumer-technology interaction in information disclosure contexts, which mostly identified factors that lead to consumer information disclosure (e.g., Kim et al., 2022; Lucas et al., 2014; Pickard et al., 2016; Pickard & Roster, 2020; Pitardi et al., 2021; Uchida et al., 2017). Different from these studies, we investigated how disclosing information influences an important company-level outcome—brand trust—as moderated by agent type. This is also a novel contribution to studies investigating brand-related outcomes of consumer's interaction with technology (e.g., Bergner et al., 2023; Cheng & Jiang, 2020; Srinivasan & Sarial-Abi, 2021), which did not examine information disclosure contexts. We show that disclosing information to AI influences an important brand-related outcome (i.e., brand trust), and this effect is substantially different from disclosing information to a human. This is an important finding as brand trust is a crucial component both for the consumers to keep engaging with the brand and for the brand to stay in the market. Research suggests that trust is an asset that is gradually established (Albert & Merunka, 2013). However, several experiments we reported in the current research show that consumers establish a level of trust using the available contextual elements in the interaction setting even without having a former relationship with the brand. Our findings indicate that disclosing information in the tech-savvy marketplace can hurt brand trust. Future studies should continue examining other downstream consequences of disclosing information to AI as companies are increasingly incorporating AI for capturing data on consumers who share their data voluntarily or involuntarily (Puntoni et al., 2021). Although disclosing information to AI impairs consumers' trust in brands, the results from Study 6 suggest that consumers enjoy the outcomes of this interaction as much as the ones who disclose information to a human. This finding implies that consumers have preconceptions about disclosing information to AI; however, this may not always lead to optimal outcomes.

We also contribute to extant knowledge on consumers' perceptions about and attitudes toward AI by documenting a novel process. The negative effect of disclosing information to AI on brand trust is driven by two novel mediators: anticipated audience size and the sense of exploitation. This novel process we identify provides new venues for future research. Certain companies or organizations might collect personal data and share it with third parties for a "better world." For instance, a web-based marketing company might collect clickthrough rate data and share it with NGOs and humanitarian organizations for them to develop effective fundraising campaigns for humanitarian causes. In such contexts, customers' inference regarding audience size may not necessarily result in a sense of exploitation. Future research should identify the boundary conditions of the effects we documented.

It is worth noting that past research has consistently found that objective or cognitively driven tasks decrease algorithm aversion as consumers make attributions and/or inferences particularly regarding the *capabilities* of algorithms for given tasks. One difference in consumers' attributions for AI and humans is that AI performs better in objective tasks (Castelo et al., 2019; Longoni & Cian, 2022; Wien & Peluso, 2021) but humans take subjective, situation- or consumer-specific information into account (Longoni et al., 2019; Zhang, 2021). In other words, these studies reveal distinct but similar processes based on consumers' inferences regarding the domain-specific capabilities of the agent. Our research complements these studies by showing that consumers judge the overall capabilities of AI in relation to humans and make inferences about them. More specifically, the agent with which consumers interact influences also the inferences regarding the number of people with which the agent may share this information. This may be driven by a potential lack of familiarity or understanding among consumers regarding the inner workings and mechanisms of AI technology. Thus, future research should explore this relationship and the role of familiarity with AI in shaping consumers' trust perceptions towards brands following information disclosure.

Note that the direct effect of agent on brand trust remained significant in Studies 2 and 4, even when controlling for the mediating roles of audience size and the sense of exploitation. This suggests a partial mediation and implies that we have unveiled one of the many potential processes that would explain why disclosing information to AI decreases brand trust compared to disclosing information to humans. Future research might uncover other potential mechanisms underlying the obtained effect.

10.2. Substantive implications

Identifying several theory-driven boundary conditions, our research offers guidance to managers on the use of algorithms in contexts where consumers are asked to disclose personal information. First, managers can prevent the detrimental effects of disclosing information to AI on brand trust by ensuring confidentiality, as confidentiality reduces the size of the inferred audience. Likewise, marketers could avoid negative consequences of information disclosure to AI by developing anthropomorphized AI agents or by deploying other techniques that will make customers attribute human-like characteristics to AI.

The documented effects are particularly pronounced among customers with higher levels of privacy concerns. This, however, does not necessarily mean that the negative brand trust-related effects of asking customers to disclose information to

an AI agent disappears for customers who are less concerned with their privacy. Thus, marketers should be aware that using AI in contexts where customers share information with the company might backfire for all customers regardless of their concern for privacy. Considering this finding, along with the finding regarding the role of information confidentiality, companies that use AI in information disclosure contexts must ensure that consumers' privacy is protected in a way that also implies a lower audience size.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research has been funded by Migros Chair Funds. The authors thank research assistants and graduate students at the Koç University Behavioral Lab.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2023.08.011.

References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751.
Albert, N., & Merunka, D. (2013). The role of brand love in consumer-brand relationships. *Journal of Consumer Marketing.*.
Anton, A. I., Earp, J. B., & Young, J. D. (2010). How internet users' privacy concerns have evolved since 2002. *IEEE Security & Privacy, 8*(1), 21–27.
Barasch, A., & Berger, J. (2014). Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research, 51*(3), 286–299.
Becerra, E. P., & Badrinarayanan, V. (2013). The influence of brand trust and brand identification on brand evangelism. *Journal of Product & Brand Management, 22*(5/6), 371–383.
Bergner, A. S., Hildebrand, C., & Häubl, G. (2023). Machine Talk: How Verbal Embodiment in Conversational AI Shapes Consumer-Brand Relationships. *Journal of Consumer Research* ucad014.
Bonezzi, A., Ostinelli, M., & Melzner, J. (2022). The human black-box: The illusion of understanding human better than algorithmic decision-making. *Journal of Experimental Psychology: General.*.
Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research, 56*(5), 809–825.
Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: The role of brand loyalty. *Journal of Marketing, 65*(2), 81–93.
Cheng, Y., & Jiang, H. (2020). How do AI-driven chatbots impact user experience? Examining gratifications, perceived privacy risk, satisfaction, loyalty, and continued use. *Journal of Broadcasting & Electronic Media, 64*(4), 592–614.
Cheng, Y., & Jiang, H. (2022). Customer–brand relationship in the era of artificial intelligence: Understanding the role of chatbot marketing efforts. *Journal of Product & Brand Management, 31*(2), 252–264.
Clark, M. S., & Waddell, B. (1985). Perceptions of exploitation in communal and exchange relationships. *Journal of Social and Personal Relationships, 2*(4), 403–418.
Coyle, J. R., Smith, T., & Platt, G. (2012). "I'm here to help": How companies' microblog responses to consumer problems influence brand perceptions. *Journal of Research in Interactive Marketing, 6*(1), 27–41.
De Cremer, D. (1999). Trust and fear of exploitation in a public goods dilemma. *Current Psychology, 18*(2), 153–163.
DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J. J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science, 23*(12), 1549–1556.
Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution, 2*(4), 265–279.
Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114.
Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research, 17*(1), 61–80.
Doney, P. M., & Cannon, J. P. (1997). An examination of the nature of trust in buyer–seller relationships. *Journal of Marketing, 61*(2), 35–51.
Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864.
Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology, 51*(4), 724–731.
Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford University Press.
Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83.
Garbarino, E., & Johnson, M. S. (1999). The different roles of satisfaction, trust, and commitment in customer relationships. *Journal of Marketing, 63*(2), 70–87.
Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior, 24*(4), 1494–1509.
Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315*(5812), 619.
Green, B., & Viljoen, S. (2020). Algorithmic realism: expanding the boundaries of algorithmic thought. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 19-31).
Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review, 10*(3), 252–264.
Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis second edition: A regression-based approach*. New York, NY: Ebook The Guilford Press.

Herbst, K. C., Finkel, E. J., Allan, D., & Fitzsimons, G. M. (2012). On the dangers of pulling a fast one: Advertisement disclaimer speed, brand trust, and purchase intention. *Journal of Consumer Research, 38*(5), 909–919.
Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication, 68*(4), 712–733.
Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM, 42*(4), 80–85.
Hollebeek, L. D., Sprott, D. E., & Brady, M. K. (2021). Rise of the machines? Customer engagement in automated service interactions. *Journal of Service Research, 24*(1), 3–8.
Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155–172.
Karataş, M., & Cutright, K. M. (2023). Thinking about God increases acceptance of artificial intelligence in decision-making. *Proceedings of the National Academy of Sciences, 120*(33) e2218961120.
Keith, M. J., Babb, J. S., Lowry, P. B., Furner, C. P., & Abdullat, A. (2015). The role of mobile-computing self-efficacy in consumer information disclosure. *Information Systems Journal, 25*(6), 637–667.
Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic interactions with a robot and robot–like agent. *Social Cognition, 26*(2), 169–181.
Kim, T., Barasz, K., & John, L. K. (2019). Why am I seeing this ad? The effect of ad transparency on ad effectiveness. *Journal of Consumer Research, 45*(5), 906–932.
Kim, T. W., Jiang, L., Duhachek, A., Lee, H., & Garvey, A. (2022). Do You Mind if I Ask You a Personal Question? How AI Service Agents Alter Consumer Self-Disclosure. *Journal of Service Research, 25*(4), 649–666.
Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology, 90*(4), 543.
Kronrod, A., Grinstein, A., & Wathieu, L. (2012). Go green! Should environmental messages be so assertive? *Journal of Marketing, 76*(1), 95–102.
Kulms, P., & Kopp, S. (2018). A social cognition perspective on human–computer trust: The effect of perceived warmth and competence on trust in decision-making with computers. *Frontiers in Digital Humanities, 14*.
Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research, 50*(5), 561–576.
Lee, D., & LaRose, R. (2011). The impact of personalized social cues of immediacy on consumers' information disclosure: A social cognitive approach. *Cyberpsychology, Behavior, and Social Networking, 14*(6), 337–343.
Lefkeli, D., Ozbay, Y., Gürhan-Canli, Z., & Eskenazi, T. (2021). Competing with or against cozmo, the robot: Influence of interaction context and outcome on mind perception. *International Journal of Social Robotics, 13*(4), 715–724.
Lin, J. S. E., & Wu, L. (2023). Examining the psychological process of developing consumer-brand relationships through strategic use of social media brand chatbots. *Computers in Human Behavior, 140* 107488.
Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650.
Longoni, C., & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The "word-of-machine" effect. *Journal of Marketing, 86*(1), 91–108.
Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37*, 94–100.
Luo, X. (2002). Trust production and privacy concerns on the Internet: A framework based on relationship marketing and social exchange theory. *Industrial Marketing Management, 31*(2), 111–118.
Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science, 38*(6), 937–947.
Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing, 81*(1), 36–58.
Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management Review, 20*(3), 709–734.
McLean, G., Osei-Frimpong, K., & Barhorst, J. (2021). Alexa, do voice assistants influence consumer brand engagement?–Examining the role of AI powered voice assistants in influencing consumer brand engagement. *Journal of Business Research, 124*, 312–328.
McLeay, F., Osburg, V. S., Yoganathan, V., & Patterson, A. (2021). Replaced by a Robot: Service Implications in the Age of the Machine. *Journal of Service Research, 24*(1), 104–121.
Mellon, J., Bailey, J., Scott, R., Breckwoldt, J., & Miori, M. (2022). Does GPT-3 know what the Most Important Issue is? *Using Large Language Models to Code Open-Text Social Survey Responses At Scale (December 22, 2022)*.
Metzger, M. J. (2004). Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *Journal of Computer-Mediated Communication, 9*(4), JCMC942.
Milne, G. R., Pettinico, G., Hajjat, F. M., & Markos, E. (2017). Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs, 51*(1), 133–161.
Minton, E. A., Kaplan, B., & Cabano, F. G. (2022). The influence of religiosity on consumers' evaluations of brands using artificial intelligence. *Psychology & Marketing, 39*(11), 2055–2071.
Morey, T., Forbath, T., & Schoop, A. (2015). Customer data: Designing for transparency and trust. *Harvard Business Review, 93*(5), 96–105.
Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing, 58*(3), 20–38.
Mothersbaugh, D. L., Foxx, W. K., Beatty, S. E., & Wang, S. (2012). Disclosure antecedents in an online service context: The role of sensitivity of information. *Journal of Service Research, 15*(1), 76–98.
Niszczota, P., & Kaszás, D. (2020). Robo-investment aversion.. *Plos One, 15*(9), e0239277.
Ormiston, M. E., Wong, E. M., & Haselhuhn, M. P. (2017). Facial-width-to-height ratio predicts perceptions of integrity in males. *Personality and Individual Differences, 105*, 40–42.
Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., & Voss, C. A. (2015). Service research priorities in a rapidly changing context. *Journal of Service Research, 18*(2), 127–159.
Pickard, M. D., Roster, C. A., & Chen, Y. (2016). Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? *Computers in Human Behavior, 65*, 23–30.
Pickard, M. D., & Roster, C. A. (2020). Using computer automated systems to conduct personal interviews: Does the mere presence of a human face inhibit disclosure? *Computers in Human Behavior, 105* 106197.
Pitardi, V., Wirtz, J., Paluch, S., & Kunz, W. (2021). Will Robots Judge Me? Examining Consumer-Service Robots Interactions in Embarrassing Service Encounters: An Abstract. In *Celebrating the Past and Future of Marketing and Discovery with Social Impact: 2021 AMS Virtual Annual Conference and World Marketing Congress* (pp. 257–258). Cham: Springer International Publishing.
Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.
Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing, 85*(1), 131–151.
Querci, I., Barbarossa, C., Romani, S., & Ricotta, F. (2022). Explaining how algorithms work reduces consumers' concerns regarding the collection of personal data and promotes AI technology adoption. *Psychology & Marketing, 39*(10), 1888–1901.
Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48*(1), 137–141.
Rifon, N. J., LaRose, R., & Choi, S. M. (2005). Your privacy is sealed: Effects of web privacy seals on trust and personal disclosures. *Journal of Consumer Affairs, 39*(2), 339–362.
Schuetzler, R. M., Grimes, G. M., Giboney, J. S., & Nunamaker Jr, J. F. (2018). The influence of conversational agents on socially desirable responding. In *Proceedings of the 51st Hawaii International Conference on System Sciences* (p. 283).
Shank, D. B., & DeSanti, A. (2018). Attributions of morality and mind to artificial intelligence after real-world moral violations. *Computers in Human Behavior, 86*, 401–411.
Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 167–196.

Srinivasan, R., & Sarial-Abi, G. (2021). When algorithms fail: Consumers' responses to brand harm crises caused by algorithm errors. *Journal of Marketing, 85*(5), 74–91.
Steenkamp, J. B. E., & Geyskens, I. (2006). How country characteristics affect the perceived value of web sites. *Journal of Marketing, 70*(3), 136–150.
Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89*(6), 845.
Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21*(3), 349–354.
Treiblmaier, H., & Chong, S. (2013). Trust and perceived risk of personal information as antecedents of online information disclosure: Results from three countries. In *Global Diffusion and Adoption of Technologies for Knowledge and Information Sharing* (pp. 341–361). IGI Global.
Uchida, T., Takahashi, H., Ban, M., Shimaya, J., Yoshikawa, Y., & Ishiguro, H. (August.). A robot counseling system—What kinds of topics do we prefer to disclose to robots? In *In 2017 26th IEEE international symposium on robot and human interactive communication (RO-MAN)* (pp. 207–212). IEEE.
Urban, G. L., Sultan, F., & Qualls, W. J. (2000). Placing trust at the center of your Internet strategy. *Sloan Management Review, 42*(1), 39–48.
Van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., & Petersen, J. A. (2017). Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences. *Journal of Service Research, 20*(1), 43–58.
Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in cognitive sciences, 14*(8), 383–388.
Weiss, A., Burgmer, P., & Mussweiler, T. (2018). Two-faced morality: Distrust promotes divergent moral standards for the self versus others. *Personality and Social Psychology Bulletin, 44*(12), 1712–1724.
Wien, A. H., & Peluso, A. M. (2021). Influence of human versus AI recommenders: The roles of product type and cognitive processes. *Journal of Business Research, 137*, 13–27.
Wiggins, J. S., & Broughton, R. (1991). A geometric taxonomy of personality scales. *European Journal of Personality, 5*(5), 343–365.
Winterich, K. P., Nenkov, G. Y., & Gonzales, G. E. (2019). Knowing what it makes: How product transformation salience increases recycling. *Journal of Marketing, 83*(4), 21–37.
Wottrich, V. M., Verlegh, P. W., & Smit, E. G. (2017). The role of customization, brand trust, and privacy concerns in advergaming. *International Journal of Advertising, 36*(1), 60–81.
Yip, J. A., & Schweitzer, M. E. (2015). Trust promotes unethical behavior: Excessive trust, opportunistic exploitation, and strategic exploitation. *Current Opinion in Psychology, 6*, 216–220.
Zhang, K. (2021). How information processing style shapes people's algorithm adoption. *Social Behavior and Personality: an International Journal, 49*(8), 1–13.
Zhang, Y., Xu, J., Jiang, Z., & Huang, S. C. (2011). Been there, done that: The impact of effort investment on goal value and consumer motivation. *Journal of Consumer., 38*(1), 78–93.
Zhang, Z., Chen, Z., & Xu, L. (2022). Artificial intelligence and moral dilemmas: Perception of ethical decision-making in AI. *Journal of Experimental Social Psychology, 101* 104327.