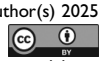
AMERICAN MARKETING ASSOCIATION

Article

Journal of Marketing
2026, Vol. 90(2) 1-26



DOI: 10.1177/00222429251344139
journals.sagepub.com/home/jmx


Automated Versus Human Agents: A Meta-Analysis of Customer Responses to Robots, Chatbots, and Algorithms and Their Contingencies

Katja Gelbrich, Holger Roschk, Sandra Miederer, and Alina Kerath

Abstract

This meta-analysis examines when three types of automated agents (AAs)—robots, chatbots, and algorithms—are equivalent to human agents (HAs) in marketing roles. An analysis of 943 effect sizes from 327 studies provides novel insights. First, customers may be skeptical of AAs; however, they value their performance and eventually choose or buy from them as if they were interacting with HAs. Second, each of the three AA types has a unique set of contingencies that affect its human equivalence; previously identified contingencies do not generalize and can even have opposing effects across AA types. This study also identifies novel contingencies, such as the multifaceted concept of artificial intelligence required to fulfill a task. Third, some contingencies make AAs more humanlike, while others make their machine characteristics salient. These findings enrich the concept of automated social presence (ASP), suggesting that AAs are hybrid beings with a social presence (i.e., the feeling of interacting with a humanlike entity) and an automated presence (i.e., the feeling of interacting with a machine entity). The authors provide recommendations on when AAs can replace HAs in marketing roles to release capacity and alleviate labor shortages. They also suggest a future research agenda.

Keywords

automated agents, artificial intelligence, robots, chatbots, algorithms, customer responses, meta-analysis, automated social presence

Submitted March 9, 2023

Automated agents (AAs) are system-based autonomous interfaces that interact with customers in the physical or digital realm to perform certain tasks (Wirtz et al. 2018).[1] They can perform human agents' (HAs) roles, such as serving coffee (Choi et al. 2023), responding to inquiries (Yalcin et al. 2022), and recommending products (Longoni and Cian 2022). Given AAs' high efficiency (Xiao and Kumar 2021), their market value is expected to reach $210.7 billion by 2033 (Saha 2023). Still, AAs lack human skills, often prompting customers to respond more negatively to them than to HAs (Longoni, Bonezzi, and Morewedge 2019). Thus, managers need to know: *When do customers view AAs as equivalent substitutes for HAs?* This will help them decide when to leverage AAs as an alternative to the default choice of human staff.

Prior studies have addressed this question, revealing contingencies that weaken the negative effect of AAs (vs. HAs) on customer responses (e.g., a utilitarian context; Longoni and

[1] The service literature often refers to service robots (Wirtz et al. 2018). We use the umbrella term automated agents and consider physically embodied robots as a subcategory.

Katja Gelbrich is Professor of Business Administration and International Management, Catholic University Eichstaett-Ingolstadt, Germany (email: Katja.Gelbrich@ku.de). Holger Roschk is Professor of Marketing, Aalborg University Business School, Aalborg University, Denmark (email: hroschk@business.aau.dk). Sandra Miederer (corresponding author) is a doctoral student of Business Administration and International Management, Catholic University Eichstaett-Ingolstadt, Germany (email: SMiederer@ku.de). Alina Kerath is a doctoral student of Business Administration and International Management, Catholic University Eichstaett-Ingolstadt, Germany (email: Alina.Kerath@ku.de).

Cian 2022). Drawing conclusions about these contingencies is difficult for two reasons. First, studies refer to different AA types: robots (physically embodied machines that interact with customers via actuators or sensors; Joshi 2022), chatbots (digital interfaces that engage customers in natural conversations; Sands et al. 2021), and algorithms (software programs that provide information by transforming an input into an output; Clegg et al. 2023). Second, studies use a variety of customer responses. Sometimes these are behaviors (e.g., choosing an AA; Holthöwer and Van Doorn 2023), but they are mostly upstream variables (e.g., trust; Wang et al. 2023). Some responses relate to the agent (e.g., perceived warmth; Frank and Otterbring 2023) and others to the firm (e.g., attitude toward the brand; Srinivasan and Sarial-Abi 2021). Hence, there is heterogeneity in which AA type is rated on which success criterion (i.e., which customer response), making it difficult to assess when AAs keep pace with HAs; this calls for consolidated research (Table 1).

The meta-analyses listed in Table 1 have consolidated variables that affect customer responses to AAs, but two gaps limit insights into their human equivalence. First, benchmarking is missing, incomplete, or often poorly suited to a marketing context. The meta-analyses on AAs alone do not use HAs as a benchmark. However, benchmarking is necessary to delineate contingencies that cause AAs to become more equivalent to HAs, and not just benefit both agents equally, yielding a null effect. For example, Blut, Wünderlich, and Brock (2024) suggest that a name can benefit AAs because it provides them with a human touch. However, customers also respond more positively to human employees using their names as part of a personalized service (Roschk and Gelbrich 2017). In contrast, the meta-analyses on AAs versus HAs use a human benchmark, which is crucial to gauge the relative effectiveness of AAs and address the substitution question. However, these studies only consider facets of AAs (e.g., face, speech) or mainly stem from communication, computer science, psychology, and journalism contexts; thus, the outcomes and contingencies do not apply to marketing.[2]

Second, no meta-analysis in Table 1 crosses the three AA types with multiple contingencies. Some of them generically focus on technology (e.g., Blut and Wang 2020) or artificial intelligence (AI) (e.g., Aguiar-Costa et al. 2022), which comprises a variety of machine applications (including AAs). Others use mixed AA samples, distinguishing between embodied (i.e., robots) and disembodied (i.e., collapsing chatbots and algorithms) AAs (Blut et al. 2021), different AI communicator roles (e.g., curator; Huang and Wang 2023), or different AI labels (e.g., AI assistants; Zehnle, Hildebrand, and Valenzuela 2025). As a result, it is difficult to assign contingencies to an AA type. Other meta-analyses examine a single AA type, but they either do not test contingencies at all (e.g., Kilani and Rajaobelina 2024) or do not assess whether the results generalize to the other two types. Generalizability is further limited in terms of customer responses. No meta-analysis tests how AAs perform across agent-related (e.g., trust in the agent) versus firm-related (e.g., attitude toward the brand) customer responses; firm-related behaviors (e.g., consumption) as such are not considered at all.

In summary, prior research cannot answer the question of when customers consider the three AA types to be equivalent to HAs in the marketing context. We conducted a comprehensive meta-analysis of multiple customer responses to robots, chatbots, and algorithms compared with human employees. Our study is based on 943 effect sizes retrieved from 327 studies with 281,954 respondents in 148 articles (Table 1). It makes three contributions to research on AAs in marketing.

First, we compare the pooled effect sizes of the three AA types (vs. HAs) for a wide range of customer responses. These are agent- or firm-related and upstream (i.e., perceptions, appraisals, intentions) or downstream (i.e., behaviors). For downstream responses, we add self-related behaviors, which occur in single studies (e.g., Desideri et al. 2019). Hence, we are the first to provide a consolidated test of customer responses with lower (i.e., agent-related, upstream) versus higher managerial relevance (i.e., firm-related, downstream). The results show that AAs tend to be less equivalent for agent-related upstream responses but equivalent for all downstream responses whether agent-, firm-, or self-related. These results challenge the belief that AAs will remain inferior to human employees in the near future (Xiao and Kumar 2021). Customers may be skeptical of AAs, but from a management perspective, this is a minor issue: Customers value AAs' performance and choose or buy from them as if they were interacting with HAs.

Second, we test contingencies for the equivalence of robots, chatbots, and algorithms (vs. HAs). The novelty of our framework lies in carving out the particularities of the three AA types and revealing idiosyncratic contingencies for their human equivalence. This overcomes the limitations of prior meta-analyses that do not distinguish between the three AA types or are restricted to one AA type, calling for systematic research on AA-type-specific contingencies. Our results put prior findings into perspective. They show that generalizing single contingencies to all AAs is undue, refuting the idea that one size fits all. We also uncover novel contingencies, such as the AI types required for a certain task (Pantano and Scarpi 2022). This disentangles the fuzziness arising from prior meta-analyses that use AI as a generic label for AAs (Aguiar-Costa et al. 2022) or intelligence as a broad contingency (Blut, Wünderlich, and Brock 2024). AI is the ability to mimic the functions of the human mind, such as problem-solving or reasoning (Longoni, Bonezzi, and Morewedge 2019). AAs may or may not be AI-enabled (Joshi 2022), and this ability has multiple facets, such as verbal-linguistic or social intelligence (Pantano and Scarpi 2022). Our study reveals tasks in which specific types of AI benefit single AA types.

[2] Other AA–HA meta-analyses are even further away from a marketing context. They stem from health care and use clinical outcomes such as arm movement after stroke (Veerbeek et al. 2017) or sleep quality in patients (Leng et al. 2019).

Table 1. This Study Fills a Void in Prior Meta-Analyses of Customer Responses to AAs.

Author (Year)	Summary of Research Question(s)[a]	AA Types[b]					Customer Responses[b]		Field	No. of Articles	No. of Effect Sizes[e]
		Generic[c]	Mixed[d]	Robots	Chatbots	Algorithms	Tests Agent- vs. Firm-Related Responses	Includes Customer Behaviors			
AAs Alone											
Aguiar-Costa et al. (2022)	How does AI adoption affect customer satisfaction in service delivery?	x							Service	19	120
Blut and Wang (2020)	What factors influence customers' technology readiness and thus their usage of new technologies?	x						AA usage	Marketing	163	2,752
Kuen et al. (2023)	How does trust in technology and its provider affect technology acceptance?	x							Service	251	657
Ma, Fan, and Mattila (2024)	What affects consumer technology adoption and experiences in hospitality?	x							Hospitality	103	471
Mehta et al. (2022)	What affects customer acceptance of AI and what theory explains this best?	x							Marketing	69	167
Blut et al. (2021)	What are the antecedents and consequences of customers anthropomorphizing AAs?		x						Marketing	71	3,404
Ladeira, Perin, and Santini (2023)	What affects the acceptance of service robots in the hospitality and tourism industry?			x					Hospitality	68	326
Blut, Wünderlich, and Brock (2024)	What affects the use of AI-based virtual assistants by retail customers?				x			AA usage	Retailing	195	2,766
Gopinath and Kasilingam (2023)	What leads customers to adopt chatbots across different service encounters?				x				Service	70	333
Li et al. (2023)	What factors affect customers' intention to adopt AI-based chatbots?				x				Service	54	153
Blut, Ghiassaleh, and Wang (2023)	Which recommendation agent type works best to support retail customers?					x			Retailing	98	480
AAs vs. HAs											
Huang and Wang (2023)	Is communication by AI agents more persuasive than that of humans?		x					Unspecified	Communication	89	300

(continued)

Table 1. (continued)

Author (Year)	Summary of Research Question(s)[a]	AA Types[b]					Customer Responses[b]		Field	No. of Articles	No. of Effect Sizes[e]
		Generic[c]	Mixed[d]	Robots	Chatbots	Algorithms	Tests Agent- vs. Firm-Related Responses	Includes Customer Behaviors			
Miller et al. (2023)	How do users respond to computer-generated versus human faces?		x						Computer science	83	925
Zehnle, Hildebrand, and Valenzuela (2025)	How context-dependent is consumer aversion to AI? Do consumer responses evolve over time? Are negative responses experimental design artifacts?		x[f]					Composite behaviors	Cross-disciplinary	72	440
Kilani and Rajaobelina (2024)	What is the impact of live chat (with a chatbot vs. human) service quality on behavioral intentions and service quality?				x[g]				Service	29	47
Stern and Mullennix (2004)	Does the persuasiveness of synthetic speech differ from human speech?				x				Psychology	7	7
Wang and Huang (2024)	How does machine vs. human authorship influence credibility perceptions and news evaluations by news recipients?					x			Journalism	26	90
Robots, Chatbots, and Algorithms vs. HAs											
This study	When do customers consider AAs to be equivalent substitutes for HAs?			x	x	x	x	Agent, firm-, and self-related	Marketing	148	943

[a]The research questions have been summarized for clarity and comparability.
[b]The categorizations of AA types and customer responses are based on the definitions used in our meta-analysis. Thus, the terms used in the sources may differ.
[c]The "generic" studies have a generic focus on AI, technology, or automation, comprising a variety of machine applications including AAs.
[d]The "mixed" studies do not distinguish between AA types but include more than one type according to our definitions.
[e]More information on the number of effect sizes is provided in Web Appendix A, Table W1.
[f]This study distinguishes between different AA types but aggregates them for the contingency analysis.
[g]This study only tests chatbots versus HAs as a moderator of one relationship in their model rather than the contingencies of an AA–HA comparison.

Table 2. The Three AA Types Have Commonalities and Particularities.

Commonalities	**All AA Types**		
Disadvantages over HAs	Lack of empathy, cognitive flexibility, and agency		
Advantages over HAs	Speed (quick data processing, instant reaction) and reliability (constant receptivity, stable results)		
Particularities	**Robots**	**Chatbots**	**Algorithms**
Definitions	Embodied machines that interact with customers through actuators or sensors (Joshi 2022).	Digital interfaces that engage customers in natural conversations (Sands et al. 2021).	Software programs that provide information to customers by transforming an input into an output (Clegg et al. 2023).
Examples	• Pepper explains hotel facilities to guests (Choi et al. 2020). • Service robot brews and serves coffee in restaurants (Choi et al. 2023).	• Replika serves as a virtual buddy (Drouin et al. 2022). • Virtual assistant gives advice in online shop (Thomaz et al. 2020).	• IBM's Watson suggests medical treatments. • Algorithm decides who gets a loan and who does not (Yalcin et al. 2022).
Mode of communication	Conversational		Informational
• Dynamics	• Dynamic		• Static
• Style	• Mimics natural human speech		• Responds to commands
Form of existence	Physical		Digital
• Body	• Tangible (interactive motor skills)		• Intangible, optionally visualized (icon or avatar)
• Space	• Real (i.e., natural environment)		• Virtual (i.e., simulated environment)

Third, we observe an overarching pattern: Some contingencies make AAs more humanlike, while others make their machine characteristics salient. These findings enrich the concept of automated social presence (ASP), an emerging paradigm in marketing (Flavián et al. 2024). ASP proposes that AAs can make customers feel that they are interacting with another social entity (Van Doorn et al. 2017). This implies that AAs' humanlike nature is valued and overlays their machine nature, which suggests that engaging customers socially is key to success. We abandon this idea and posit that AAs are hybrid beings with a social presence (i.e., the feeling of interacting with a humanlike entity) *and* an automated presence (i.e., the feeling of interacting with a machine entity). Social presence only clearly benefits one AA type; it can be irrelevant to or even harm others. Automated presence benefits all AA types, which can leverage both advantageous and presumed disadvantageous machine characteristics. Hence, AAs do not need to be humanlike but can be equivalent substitutes for HAs just because they are perceived as machines.

The structure of this article follows a guided test of AA-type-specific contingencies. We describe the properties of robots, chatbots, and algorithms; use them to develop hypotheses; test these hypotheses; and use the results to inform theory. Since AA research lacks an overarching theory (De Keyser and Kunz 2022), this approach is suitable. A guided test, rather than an empirics-first approach (Golder et al. 2023), is chosen because the three AA types' particularities allow for systematic reasoning on specific contingencies for them. We use the results to predict when AAs will keep pace with HAs. This helps marketers decide when to consider robots, chatbots, or algorithms as alternatives to HAs, or when it is not advisable to do so. We also provide a research agenda to further develop ASP.

Conceptual Development

Properties of Robots, Chatbots, and Algorithms

All AA types have a common technological nature with disadvantages and advantages over HAs (Table 2). Disadvantages relate to affective abilities: AAs (vs. HAs) lack empathy; they cannot experience emotions and often do not recognize them (Zhou et al. 2023). AAs also fall short in their cognitive abilities, particularly in cognitive flexibility (Longoni, Bonezzi, and Morewedge 2019), as they cannot think outside the box due to their reliance on past interactions (Wirtz et al. 2018). They also lack agency in that they have no opinions, cannot judge (i.e., cannot distinguish right from wrong), and do not feel responsible (Holthöwer and Van Doorn 2023). However, AAs have cognitive advantages related to speed and reliability. They can quickly process big data, react instantly, are always receptive, and replicate stable results (Larkin, Drummond Otten, and Árvai 2022).

All AA types also carry particularities that relate to their mode of communication and form of existence (Table 2). AAs' mode of communication can be conversational, characterized by dynamic mimicry of natural human speech. This applies to robots, which can simulate highly interactive conversations with customers (Pitardi et al. 2022). These are mostly voice-

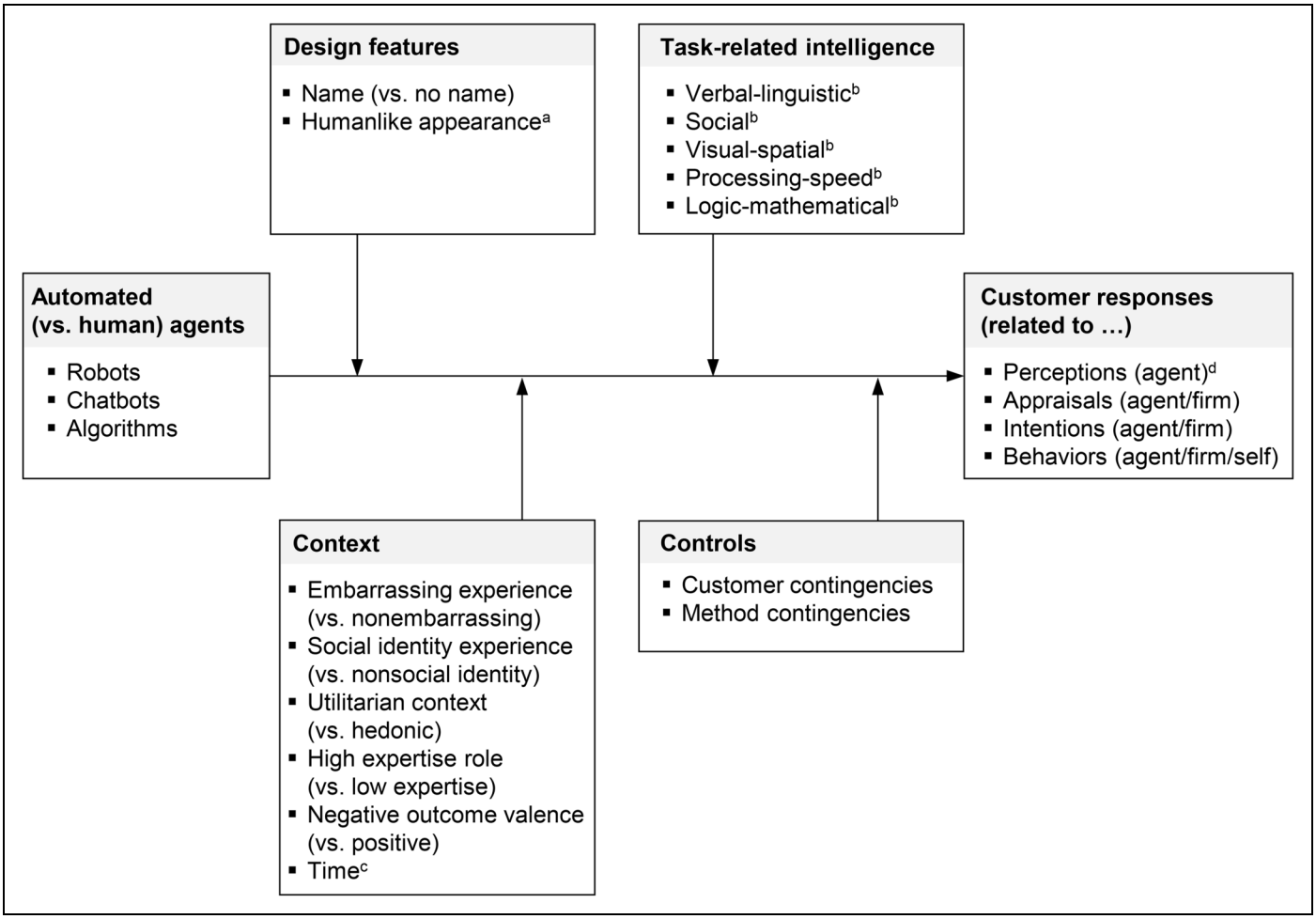


Figure 1. Meta-Analytic Framework for Customer Responses to AAs Versus HAs.
[a]"Very low" to "very high."
[b]"Not at all required" to "very much required."
[c]Year 2000 to 2023.
[d]This category includes three variables: perceived humanlikeness, warmth, and competence of the agent.

based but can also be text-based (i.e., holding a screen) (Castelo et al. 2023). Chatbots' mode of communication also simulates highly interactive human speech (Luo et al. 2019), which is mostly text-based (Ruan and Mezei 2022) but can also be voice-based (Ruiz-Equihua et al. 2023). The communication mode can also be informational, characterized by static responses to commands. This applies to algorithms that provide information but do not engage in sophisticated conversations. They transform a one-time input into a requested output (Clegg et al. 2023) via sophisticated data-processing techniques (Schoeffer, Machowski, and Kuehl 2021).

AAs' form of existence can be physical, with a tangible body whose motor skills enable interactions with customers in real space. This applies to robots; their materialized bodies are imbued with electronic components (Joshi 2022), allowing for skillful interactions with customers in their natural environment (Čaić et al. 2020). Robots can touch, move, and process items for customers (e.g., clean a car) or perform services on customers' bodies (e.g., cut hair) (Wirtz et al. 2018). AAs can also have a digital form of existence, with an intangible "body" that can be visualized with icons or avatars. This applies to chatbots (Crolic et al. 2022) and algorithms (Wien and Peluso 2021). Such digital agents exist only in virtual space.

Conceptual Framework

Figure 1 depicts our conceptual framework. The independent variable refers to automated (vs. human) agents. Given the idiosyncratic properties of each AA type, we compare robots, chatbots, and algorithms with HAs separately. As a baseline, we examine the pooled effect size of these three AA types (vs. HAs) on customer responses as the dependent variable. The customer responses from single studies are organized into four categories. One category captures the perceptions of the agent. These are perceived humanlikeness—the degree to which the agent is ascribed humanlike emotions, characteristics, or

motivations (Epley, Waytz, and Cacioppo 2007); perceived warmth—the degree to which the agent is perceived as having positive intentions; and perceived competence—the degree to which the agent is perceived as being capable (Fiske, Cuddy, and Glick 2007). The other three categories are responses from the general marketing literature: customer appraisals, intentions, and behaviors (Hoyer, MacInnis, and Pieters 2024). They can relate to the agent or firm (Ryoo, Jeon, and Kim 2024), for example, rapport with the agent or satisfaction with the firm as agent- or firm-related customer appraisals. For behaviors, our framework also covers self-related responses because they occur in several studies (e.g., customers' task accuracy; Desideri et al. 2019). Web Appendix B provides details.

Our framework includes four categories of contingencies for the assumed negative effect of the three AA types (vs. HAs) on customer responses. Three categories contain substantive contingencies. The choice of these categories is guided by their alleged ability to affect the human equivalence of AAs. We include design features, which relate to the external characteristics of AAs. Design features may be relevant for the equivalence question because they represent easily accessible cues that may make AAs appear more like HAs (Go and Sundar 2019). We include task-related intelligence, defined as the degree to which AAs require a specific form of AI to fulfill a task (Pantano and Scarpi 2022). This contingency category may be relevant because it refers to internal characteristics of AAs that may make them akin to HAs, which become accessible to customers during interactions. We also consider context, which refers to the circumstances of an interaction. It is included because AAs' (dis)advantages over HAs may matter in certain circumstances, making AAs perform as well as, or even worse than, HAs. Control contingencies represent the fourth category, with variables regarding customers and methods. Controls are included as covariates without setting up hypotheses for them (Melnyk, Carrillat, and Melnyk 2022).

Contingencies for the Human Equivalence of AAs

Overview

For the substantive contingencies (i.e., design features, task-related intelligence, and context), we follow a guided test of hypotheses. The idea is that these contingencies are AA-type-specific because each AA type has a certain mode of communication (conversational vs. informational) and form of existence (physical vs. digital) (Table 2). Thus, we first justify the choice of a contingency by explaining why it is relevant for AAs with a specific mode of communication or form of existence. Then, we argue why we expect this contingency to increase or further decrease the human equivalence of these very AAs—that is, whether it may weaken or further strengthen their presumed negative effect on customer responses.

Design Features

A *name (vs. no name)* refers to whether the AAs are assigned a distinctive designation (El Halabi and Trendel 2024). This variable may matter for AAs with a conversational mode of communication (i.e., robots and chatbots; Table 2). This is because conversational AAs are made for highly interactive exchanges in which information between two parties is exchanged (Luo et al. 2019). Hence, customers need to address and engage with these AAs, as in human-to-human talk. We expect that a name will help them do this by providing their machine interlocutor with an identity label (Go and Sundar 2019). This label is likely to facilitate conversations with AAs and make customers feel connected to them (McLean, Osei-Frimpong, and Barhorst 2021). This may be even more true given that conversational AAs are also able to address customers by name during dialogs (Čaić et al. 2020). Thus, we hypothesize that a name will increase the human equivalence of conversational AAs:

> **H_1:** The negative effect of (a) robots and (b) chatbots (relative to HAs) on customer responses weakens when they have (vs. do not have) a name.

A *humanlike appearance* is the degree to which AAs look humanlike (Mende et al. 2019). This variable may matter for AAs with a physical form of existence (i.e., robots; Table 2) because physical AAs' tangible bodies make their appearance a conspicuous quality. We expect that a visual cue, such as a humanlike appearance, will dampen the perception of physical AAs as machines (Dootson et al. 2023). Making their bodies more humanlike may lead customers to attribute human qualities to them, such as more personalized treatment and a better experience (Belanche et al. 2021), thereby improving the quality of the interaction (Wirtz et al. 2018). In addition, customers may associate physical AAs with a humanlike appearance with cutting-edge technology (Belanche et al. 2021). Prior research has also questioned these positive effects, suggesting that humanoid physical AAs may appear creepy (Mende et al. 2019). However, this phenomenon, known as the uncanny valley hypothesis (Mori, MacDorman, and Kageki 2012), refers to near-perfect humanoid machine bodies (Wirtz et al. 2018) and could not to be universally validated by consolidating research (Kätsyri et al. 2015). Thus, in acknowledging an overall positive tendency, we hypothesize that a humanlike appearance will increase the human equivalence of physical AAs:

> **H_2:** The negative effect of robots (relative to HAs) on customer responses weakens as their humanlike appearance increases.

Task-Related Intelligence

Regarding the form of AI required to fulfill a task, different taxonomies exist. Huang and Rust (2018) posit that AAs can reach successive intelligence levels during their development.

Pantano and Scarpi (2022) suggest different types of AI: verbal-linguistic, social, visual-spatial, processing-speed, and logic-mathematical intelligence. These AI types are more instrumental because they allow for a more effective mapping of the AAs to the tasks they have to perform. Therefore, we use this taxonomy.

Verbal-linguistic intelligence is the degree to which a task requires the ability to understand and mimic human language (Pantano and Scarpi 2022). This variable may matter for AAs with a conversational mode of communication (i.e., robots and chatbots; Table 2) because conversational AAs engage customers in natural, one-on-one talk (Luo et al. 2019). As such, customers should be able to assess the verbal skills of these AAs during interactions. We expect that conversational AAs will benefit from their talent in this regard. Their interfaces rely on speech recognition and language processing (Pitardi et al. 2022). They are designed for the very purpose of understanding and mimicking human language (Luo et al. 2019). Like humans, conversational AAs can use idioms and phrases of politeness (Pitardi et al. 2022) as well as emojis in written conversations (Beattie, Edwards, and Edwards 2020). Hence, the more a task requires this type of intelligence, such as explaining nutritional details of food, the more conversational AAs can leverage their linguistic skills. Prior research shows that hotel guests feel attracted by conversational AAs mimicking verbal human demeanor, such as proactively greeting guests and asking if they need help (Pan et al. 2015). Similarly, conversational AAs that apologize for a long wait at a hotel check-in are perceived as sincere (Zhang et al. 2023). Thus, we hypothesize that increasing requirements for verbal-linguistic intelligence will increase the human equivalence of conversational AAs:

H_3: The negative effect of (a) robots and (b) chatbots (relative to HAs) on customer responses weakens as the required verbal-linguistic intelligence increases.

Social intelligence is the degree to which a task requires the ability to interact with humans; it includes understanding emotions and responding to social cues and is associated with empathy (Pantano and Scarpi 2022). Social intelligence may matter for AAs with a conversational mode of communication (i.e., robots and chatbots; Table 2). This is because employee–customer talks are highly interactive exchanges that go beyond verbal intercourse and incorporate a social component: Customers can gauge whether employees display the expected empathetic concern for their needs (Wieseke, Geigenmüller, and Kraus 2012). We expect conversational AAs to suffer from AAs' shortcomings in this regard, particularly from their lack of empathy (Zhou et al. 2023; Table 2). Even for conversational AAs, this ability is difficult to imitate, supporting customers' prejudice that even verbally fluent AAs still lack empathy (Luo et al. 2019). Even more, they may not believe that machines can feel at all, and perceive conversational AAs that show emotions as creepy (Appel et al. 2020). Hence, the more a task requires empathy, such as discussing customers' health, the more conversational AAs' lack of empathy may become salient. Thus, we hypothesize that increasing requirements for social intelligence will decrease the human equivalence of conversational AAs:

H_4: The negative effect of (a) robots and (b) chatbots (relative to HAs) on customer responses is strengthened as the required social intelligence increases.

In our context, *visual-spatial intelligence* can be defined as the degree to which a task requires the ability to understand space and identify patterns (Pantano and Scarpi 2022). This variable may matter for AAs with a physical form of existence (i.e., robots; Table 2). This is because physical AAs have tangible bodies with interactive motor skills that move in real space. This gives customers a touch-and-feel experience of whether these agents are moving skillfully in their environment. We expect that physical AAs are likely to benefit from AAs' outstanding abilities in this regard, namely their speed and reliability (Larkin, Drummond Otten, and Árvai 2022; Table 2). They allow physical AAs to move skillfully (Sanders et al. 2019) and enable them to quickly identify obstacles by scanning their surfaces and dimensions to navigate around them (Čaić et al. 2020). Using sensors, physical AAs are increasingly able to assess spatial proximity and respond to stimuli, making them almost as dexterous as humans (Joshi 2022). Hence, the more a task requires visual-spatial intelligence, such as giving directions at the airport, the more physical AAs may leverage their speed and reliability. Accordingly, we hypothesize that increasing requirements for visual-spatial intelligence will increase the human equivalence of physical AAs:

H_5: The negative effect of robots (relative to HAs) on customer responses weakens as the required visual-spatial intelligence increases.

Processing-speed intelligence is the degree to which a task requires the ability to perform repetitive functions rapidly and fluently (Pantano and Scarpi 2022). This variable may matter for AAs with a physical form of existence (i.e., robots; Table 2). Again, this is because these AAs are embodied machines operating in real space. Unlike in digital space, customers get a touch-and-feel experience. They can see and feel the embodied AAs moving around, and gauge whether they are performing repetitive jobs quickly and accurately. We expect that physical AAs will benefit from AAs' advantageous speed and reliability (Table 2) in tasks that directly refer to these abilities. Examples are retrieving items from a warehouse (Sanders et al. 2019) or bringing guests' luggage to their rooms (Wirtz et al. 2018). Therefore, the more a task requires processing-speed intelligence, the more physical AAs can demonstrate their superior speed and reliability. We hypothesize that increasing requirements for processing-speed intelligence will increase the human equivalence of physical AAs:

H_6: The negative effect of robots (relative to HAs) on customer responses weakens as the required processing-speed intelligence increases.

Logic-mathematical intelligence is the degree to which the task performed by an agent requires the ability to logically analyze situations and find solutions (Pantano and Scarpi 2022). This variable may matter for AAs with an informational mode of communication (i.e., algorithms; Table 2). These AAs take a command, analyze data, and transform it into an output (Table 2). Logical-analytical skills are crucial for this process (Clegg et al. 2023). We expect that informational AAs will benefit from machine agents' superior abilities in this regard, namely from their speed and reliability (Table 2). Both skills help them succeed at "number crunching." Informational AAs run complex calculations with dependable output, helping customers make informed decisions (Schoeffer, Machowski, and Kuehl 2021). Hence, the more a task requires logic-mathematical intelligence, such as providing financial advice for a retirement portfolio, the more informational AAs may be able to demonstrate their speed and reliability. Accordingly, we hypothesize that increasing requirements for logic-mathematical intelligence will increase the human equivalence of informational AAs:

H_7: The negative effect of algorithms (relative to HAs) on customer responses weakens as the required logic-mathematical intelligence increases.

Context

An *embarrassing (vs. nonembarrassing) experience* makes customers feel uncomfortable because they fear that others will judge them negatively (Dahl, Manchanda, and Argo 2001). This variable may matter for AAs with a conversational mode of communication (i.e., robots and chatbots; Table 2). Conversational AAs are made for highly interactive settings where customers tend to expose themselves to their conversation partner (Jin, Walker, and Reczek 2024)—an aspect that holds even more true because customers cannot prepare for unexpected questions in dynamic exchanges (Holthöwer and Van Doorn 2023). We expect that in embarrassing situations, conversational AAs will benefit from one of the alleged disadvantages of AAs: their lack of agency (Table 2). Since customers believe that machines cannot form their own opinions (Pitardi et al. 2022), they trust that such mindless entities will not judge them for buying embarrassing products, which reduces shame (Holthöwer and Van Doorn 2023). Hence, in awkward situations, such as when buying an antifungal treatment, conversational AAs can leverage their lack of agency. Thus, we hypothesize that an embarrassing experience will increase the human equivalence of conversational AAs:

H_8: The negative effect of (a) robots and (b) chatbots (relative to HAs) on customer responses weakens for an embarrassing (vs. nonembarrassing) experience.

A *social identity (vs. nonsocial identity) experience* refers to whether the interaction stimulates customers' wish for self-improvement, driven by the desire to be perceived positively by others (Schmitt 1999). This variable may matter for AAs with a physical form of existence (i.e., robots; Table 2). Physical AAs interact with customers in real space. These natural environments are often public settings (Mende et al. 2019), where others can observe interactions (e.g., a classroom). This public exposure may be relevant when customers seek to self-improve because it affects their image (Allard and White 2015). We expect that physical AAs will suffer in these contexts because of AAs' general lack of empathy (Pitardi et al. 2022; Table 2). Customers seeking self-improvement work on their ideal self (Schmitt 1999); they may expect compassionate feedback from the agents that are helping them overcome their deficits. In public settings, nonempathetic feedback may harm customers' image and denigrate them in front of bystanders. Hence, in contexts involving a social identity experience (e.g., when learning a new skill; Li et al. 2016), physical AAs may suffer from their lack of empathy. Accordingly, we hypothesize that a social-identity experience will decrease the human equivalence of conversational-physical AAs:

H_9: The negative effect of robots (relative to HAs) on customer responses is strengthened for a social identity (vs. nonsocial identity) experience.

Utilitarian (vs. hedonic) context refers to cognitive decisions based on functional benefits; a hedonic context involves affective decisions based on experiential pleasure (Khan, Dhar, and Wertenbroch 2005). This variable may matter for AAs with an informational mode of communication (i.e., algorithms; Table 2). The function of informational AAs is to provide output (Table 2) that helps customers make decisions (Schoeffer, Machowski, and Kuehl 2021). Thus, whether a decision is cognitively or affectively driven should play a role in this type of AA. We expect that informational AAs will benefit in utilitarian contexts because of the speed and reliability of AAs (Table 2). In utilitarian contexts, customers' decisions are driven by rationality (Khan, Dhar, and Wertenbroch 2005), which helps them achieve instrumental goals such as healthy nutrition (Longoni and Cian 2022) or financial yields (Larkin, Drummond Otten, and Árvai 2022). Hence, receiving a quick, reliable output may be appreciated. In contrast, informational AAs may perform poorly in hedonic contexts, such as recommending a concert. Here, customers expect experiential pleasure (Longoni and Cian 2022); speed and reliability seem less crucial. Thus, we hypothesize that a utilitarian (vs. hedonic) context will increase the human equivalence of informational AAs:

H_{10}: The negative effect of algorithms (relative to HAs) on customer responses weakens in a utilitarian (vs. hedonic) context.

A *high-expertise (vs. low-expertise) role* refers to AAs replacing HAs in roles that require a rather high (vs. low) level of qualification (Xie et al. 2022) and knowledge (Önkal et al. 2009). This context variable may matter for AAs with an informational mode of communication (i.e., algorithms; Table 2). These AAs provide requested outputs (Clegg et al. 2023), raising the question of whether this static communication meets the requirements of highly qualified agents. We expect that this AA type will suffer in high-expertise roles because these roles make one of AAs' disadvantages salient: their lack of cognitive flexibility (Table 2). AAs perform generally badly at thinking outside the box (Wirtz et al. 2018), but informational AAs' rigid communication style is particularly limited. In contrast to the dynamic communication style of conversational AAs, they only transform a one-time input into a requested output. Hence, informational AAs are unable to be creative, improvise, or explain unforeseen results to customers. Since these are typical requirements of experts (Önkal et al. 2009; Zhang, Pentina, and Fan 2021), informational AAs may suffer from their lack of cognitive flexibility in high-expertise roles (e.g., doctors). In contrast, customers may accept informational AAs' static communication in low-expertise roles (e.g., recommending the right clothing sizes). Thus, we hypothesize that a high-expertise (vs. low-expertise) role will decrease the human equivalence of informational AAs:

H_{11}: The negative effect of algorithms (relative to HAs) on customer responses is strengthened for a high-expertise (vs. low-expertise) role.

The two remaining context contingencies from our framework (Figure 1) are included as empirical examinations because they may affect the human equivalence of all AA types. A *negative (vs. positive) outcome valence* refers to whether the result of an interaction is unfavorable (vs. favorable), including worse-than-expected outcomes (Garvey, Kim, and Duhachek 2022), service failures (Srinivasan and Sarial-Abi 2021), and denial of service (Yalcin et al. 2022). This variable is included because customers tend to engage in attribution processes about the causer of negative events and evaluate mindless machines differently than humans (Srinivasan and Sarial-Abi 2021). We expect that a negative outcome valence will increase the human equivalence of AAs due to their lack of cognitive flexibility (Table 2). This may make customers feel that AAs follow standardized procedures (Yu, Xiong, and Shen 2022) and neglect their uniqueness (Longoni, Bonezzi, and Morewedge 2019). Thus, customers tend to take a negative outcome (e.g., being denied a loan) from AAs less personally than from human employees (Yalcin et al. 2022). This mechanism may apply to all AA types, as it is easier to accept an unfavorable outcome from any kind of machine.

Time refers to the studies' chronological age. We included this variable because technology changes over time. We expect that AAs will move toward equivalence to HAs over time because customers are increasingly exposed to AAs in their daily lives (De Keyser and Kunz 2022) and may get used to them. Furthermore, AAs should become more capable over time because newer AAs are connected to cloud-based systems (Wirtz et al. 2018) and imbued with AI (Huang and Rust 2021). We assume a positive effect for all AA types because increased exposure and technological improvements apply to all of them.

Nonlinear Effects

Nonlinear effects are conceivable for continuous contingencies, such as humanlike appearance, task-related intelligence, and time. We focus on humanlike appearance because it is the only contingency with a theoretical rationale for potential nonlinear effects. Prior research suggests the possibility of an uncanny valley for physical AAs (i.e., robots) with an almost humanlike appearance (Mori, MacDorman, and Kageki 2012). Hence, in addition to the proposed positive, linear effect of a humanlike appearance for this AA type (H_2), we explore nonlinear patterns. While uncanny valley research focuses on physical AAs (Kim, De Visser, and Phillips 2022), we also explore nonlinear effects for digital AAs (i.e., chatbots and algorithms) because they can be visualized.

Method

Data Collection

Literature search. We used five search strategies to identify relevant studies on the effects of AAs (vs. HAs) on customer responses. First, we searched electronic databases (EBSCO, ScienceDirect, and Web of Science) for articles in relevant fields (e.g., marketing, psychology, and computer science) using search terms such as robot, chatbot, algorithm digital agent, voice assistant, and AI in combination with replace/comparison and human/employee. Second, we manually reviewed relevant journals in the fields of marketing, service, psychology, and human–machine interaction (e.g., *Computers in Human Behavior*, *Journal of Consumer Psychology*, *Journal of Consumer Research*, *Journal of Marketing*, *Journal of Marketing Research*, *Journal of Service Research*, *Psychological Science*) covering the period from January 2000 (when the first studies appeared) through March 2022; a second search covered the period through June 2023 to update the dataset. Third, we searched the internet (e.g., SSRN, PsyArXiv, Google Scholar) for gray literature (i.e., dissertations, preprints, and conference proceedings) to minimize publication bias (Harrer et al. 2021). Fourth, we used the retrieved articles for a backward citation search and Google Scholar for a forward citation search. Fifth, we screened the references available at the time of our study from related meta-analyses in Table 1. Throughout the process, we asked the authors of the

identified articles for additional data to retrieve effect sizes (when necessary) and any unpublished work.

Four inclusion criteria were applied. First, the studies had to compare an AA with an HA. Second, they had to address an employee–customer interaction (e.g., selling products, providing services). To account for related fields, we also included implicit employee–customer interactions in which the agents' demeanor could be associated with a marketing context (e.g., people making decisions based on given information or chatting for social support). Third, the studies had to use at least one of the customer responses in our framework (Figure 1). Fourth, we included only studies for which we could extract or calculate the effect size. As exclusion criteria, we omitted studies (1) in which an AA supported rather than replaced an HA or in which an AA replaced an employee's colleague, (2) that referred to contexts too distant from marketing (e.g., preschools, mine clearance), or (3) that compared HAs with self-service technologies (e.g., ATMs), given that they are not AAs owing to their low level of autonomy (Holthöwer and Van Doorn 2023).

Database. The final database consisted of 148 articles, including 327 independent experimental studies[3] published from January 2000 to June 2023. Overall, 29.7% of the studies were real interactions (vs. scenarios); of these, 67.0% were field (vs. lab) experiments. Participants' mean age was 32.8 years, and the average proportion of female participants was 48.9%. In the studies that specified their sample, participants were from North America (47.9%), Europe (15.6%), Asia (33.8%), and other regions (2.5%). Web Appendices C and D list the included studies and articles, respectively. We excluded two outliers with values more than three times the standard deviation (Liadeli, Sotgiu, and Verlegh 2023), resulting in a dataset of 943 effect sizes. The results of the hypothesis testing, as well as the relative comparisons of effect sizes across AA types and customer responses, remained stable when these outliers were included.

Calculation of effect sizes. We first calculated the standardized mean differences (Cohen's d) between the AA and HA groups. When the means and SDs were not given, we calculated Cohen's d from the reported statistics (e.g., χ^2, F, t, *p*-values) via Lipsey and Wilson's (2001) formulas. Next, we converted d into the Pearson correlation coefficient r (Lipsey and Wilson 2001), which is a widely used metric for meta-analyses in marketing (e.g., Melnyk, Carrillat, and Melnyk 2022). A positive (negative) r value indicates that AAs generate more positive (negative) customer responses than HAs do.

[3] Twelve experiments can be considered quasi-experiments, in which participants were assigned to the AA (vs. HA) group based on their prior experience or customers' preferences (e.g., Ivanov, Webster, and Seyitoğlu 2023; Prentice and Nguyen 2020).

Coding scheme. Table 3 shows the coding scheme for the contingencies from our framework (Figure 1). Two independent nonauthors, unaware of the hypotheses, used this scheme to code the contingencies (Melnyk, Carrillat, and Melnyk 2022) and the three AA types (robots, chatbots, and algorithms) based on their definitions. Most variables were subjective and required high-inference coding, using the operationalizations rather than the labels in the studies. We added examples of scale values from the studies to the coding scheme to increase the validity of the coding (Cooper 2017). Furthermore, most variables were dichotomous. For the continuous contingencies humanlike appearance and the intelligence types, we used a 1–9 scale. Krippendorff's alpha values for intercoder agreement exceeded the .60 threshold, with 12 of 17 codings exceeding .80, indicating substantial agreement. Inconsistencies were discussed with the research team until a solution was found. Objective contingencies were coded by the authors: time, percentage of female participants, mean age, control variables, published work, and top-tier journal.

Data Analysis

Pooling effect sizes. The extracted effect sizes r were corrected for reliability to account for attenuation from random measurement error, r_c (Hunter and Schmidt 2004). Missing reliability values were replaced by the mean reliability for each variable across studies. We performed a Fisher's z transformation for r_c to reduce range restrictions and to account for biased standard error estimates in small samples (Alexander, Scozzaro, and Borodkin 1989). Next, we calculated the pooled effect size $\bar{r}$ for each customer response across robots, chatbots, algorithms, and AAs overall. Here, we weighted effect sizes by their inverse variance, giving greater weight to more precise effects. We used a random-effects model to account for expected heterogeneity in the effect sizes and a multilevel approach because the effect sizes are nested within studies (Harrer et al. 2021). Web Appendix F shows the formulas for these calculations. Models were estimated with the metafor package for R and the rma.mv function (Viechtbauer 2010).

Meta-regressions. The contingencies were tested via meta-regressions with the same weighting, random effects, and multilevel specifications used for pooling effect sizes. This approach was justified by the I^2 values for robots, chatbots, algorithms, and AAs overall (see Web Appendix G), indicating a high percentage of effect size variability that cannot be explained by mere sampling error (Higgins and Thompson 2002). We performed multiple meta-regressions, regressing the Fisher's z-transformed effect sizes on the contingency variables. Specifically, we estimated a robot ($k = 264$), a chatbot ($k = 261$), and an algorithm ($k = 418$) model in the respective subsamples. The models accounted for the customer responses (see Web Appendix B) with dummy variables and included the substantive and control contingencies from our framework (Table 3). We also estimated an AAs overall model ($k = 943$)

Table 3. Coding Scheme and Statistical Properties of Contingencies.

Variable	Coding Scheme	Mean (SD)
Substantive Contingencies		
Design feature		
Name	Does the AA have a name (1) or not (0)?	.29 (.45)
Humanlike appearance[a]	How humanlike does the AA look? (1 = "not at all," and 9 = "very much")	4.48 (1.81)
Task-related intelligence		
Verbal-linguistic	Does the task performed by the agent require the ability to understand and mimic human language? (1 = "not at all," and 9 = "very much")	4.08 (1.94)
Social	Does the task performed by the agent require the ability to interact with humans, including understanding emotions and responding to social cues? (1 = "not at all," and 9 = "very much")	3.08 (1.95)
Visual-spatial	Does the task performed by the agent require the ability to understand space and identify patterns? (1 = "not at all," 9 = "very much")	2.34 (2.31)
Processing-speed	Does the task performed by the agent require the ability to perform repetitive jobs rapidly and fluently? (1 = "not at all," 9 = "very much")	5.36 (1.53)
Logic-mathematical	Does the task performed by the agent require the ability to logically analyze situations or problems and find solutions? (1 = "not at all," 9 = "very much")	5.24 (1.85)
Context		
Embarrassing experience	Is the context embarrassing for customers (1) or not (0)?	.09 (.28)
Social identity experience	Does the experience stimulate customers' wish for self-improvement (1) or not (0)?	.14 (.35)
Utilitarian context	Is the context of the task utilitarian (i.e., reflecting functional benefits) (1) or hedonic (i.e., reflecting affective benefits) (0)?	.62 (.48)
High-expertise role	Does the agent take a high-expertise role (i.e., requiring higher levels of qualification) (1) or a low-expertise role (i.e., requiring lower levels of qualification) (0)?	.24 (.42)
Negative outcome valence	Is the outcome of the interaction unfavorable (1) or favorable (0) for the customer?	.13 (.32)
Time[b]	Publication year	2020 (4.66)
Control Contingencies		
Customer		
Proportion of female participants	Proportion of female participants in sample	.49 (.19)
Mean age	Mean age of participants (in years)	32.77 (7.32)
Method		
Control variables	Did the analyses include control variables (1) or not (0)?	.07 (.26)
Published work	Was the article published (1) or unpublished (0)?	.91 (.29)
Top-tier[c]	Was the article published in a top-tier journal (1) or not (0)?	.21 (.40)
General population	Were the participants from the general population (1) or students (0)?	.79 (.41)
Crowd-based recruitment	Was the recruitment of participants crowd-based (1) or non-crowed-based (0)?	.67 (.47)
Business journal	Was the article published in a business journal (1) or not (0)?	.62 (.49)
Real interaction[d]	Was the research design based on a real field experiment (1/3), a real lab experiment (1/3), or a scenario (−2/3)?	−.33 (.47)
Field experiment[d]	Was the research design based on a real field experiment (1/2), a real lab experiment (−1/2), or a scenario (0)?	.04 (.28)

[a]A humanlike appearance differs from anthropomorphism (Blut et al. 2021), which is not an objective design feature but rather a tendency to imbue nonhuman agents with humanlike characteristics (Epley, Waytz, and Cacioppo 2007). To ensure an objective assessment, we applied the humanlike appearance scale to all AA types, as chatbots and algorithms (if visualized) can be avatars with full bodies and sometimes look like physical robots (examples in Web Appendix E).
[b]We used publication year as a proxy variable for the year of data collection, since most studies did not report this information.
[c]Because our dataset included studies from multiple disciplines, we relied on the journal impact factor (JIF) as an available cross-disciplinary indicator of journal rank, using JIF ≥ 10 as a threshold for top-tier (vs. non-top-tier) journals.
[d]These two variables were contrast-coded because a field versus lab experiment is nested in real interaction. Specifically, the variable real interaction indicates whether an experiment was any kind of real interaction (1/3) or a scenario (−2/3). The variable field experiments specifies the real interactions as field (1/2) versus lab (−1/2) experiments.

with the full dataset, accounting for the AA types (robots, chatbots, and algorithms) with two dummy variables. Web Appendix H provides the models' formulas.

Multicollinearity. In the meta-regressions, we accounted for multicollinearity by removing contingencies with high variance inflation factors (VIFs) in two steps. First, to obtain a

consistent set of control contingencies, we removed general population and crowd-based recruitment from all models (as they had VIFs ≥ 3.0). In addition, we had to remove time from the chatbot and algorithm models because it also caused VIFs ≥ 3.0 in the control contingencies. We tested the removed contingencies separately in ex post analyses. Second, we removed nonsignificant substantive contingencies with VIFs ≥ 3.0 one by one, starting with the highest VIF, until all VIFs were < 3.0. The resulting maximum VIFs were then 2.84 (robots), 2.96 (chatbots), 2.79 (algorithms), and 2.87 (AAs overall), which are below those reported in other meta-analyses (e.g., Melnyk, Carrillat, and Melnyk 2022). Likelihood ratio tests (Harrer et al. 2021) revealed that removing variables with high VIFs did not reduce model fit (robots: $p = .982$, chatbots: $p = .435$, algorithms: $p = .520$). The removed variables were social intelligence and high-expertise role (robots), visual-spatial intelligence (chatbots), and verbal-linguistic intelligence (algorithms). We tested those variables in ex post analyses, retaining the formerly removed substantive contingencies and dropping those with the next-highest VIF until all VIFs were < 3.0. Web Appendix I presents details and correlation tables.

Publication bias. We checked for publication bias (i.e., unpublished nonsignificant effects; Harrer et al. 2021). The fail-safe N-values for the pooled effect sizes and the funnel plots did not provide evidence of publication bias. We also included effect size precision as an independent variable in our meta-regression models (Melnyk, Carrillat, and Melnyk 2022). It had no effect in the robot (b = .001, $p = .692$), chatbot (b = −.001, $p = .805$), algorithm (b = .001, $p = .513$), or AAs overall (b = .002, $p = .214$) models, further indicating that publication bias is not a severe problem. Web Appendix J provides details.

Results

Pooled Effect Sizes

Table 4 depicts the pooled effect sizes for the AA types (i.e., robots, chatbots, algorithms, and AAs overall) and customer responses (i.e., perceptions, appraisals, intentions, behaviors, and responses overall). The overall pooled effect size $\bar{r}$ (i.e., AAs and responses overall) is −.127 ($p < .001$), which is a small negative effect according to Cohen (1988). For AAs overall, the pooled effects sizes for perceptions, appraisals, and intentions range between −.380 ($p < .001$) and −.074 ($p < .001$). Those for behaviors range between −.038 ($p = .680$) and −.014 ($p = .788$).

A row-wise comparison of the AA types reveals some differences (Table 4, last column). For the responses overall, the negative effect size is larger for robots ($\bar{r} = -.167$) than for the two digital agents (algorithms: $\bar{r} = -.105$ and chatbots: $\bar{r} = -.122$; $p = .042$). Robots' negative effect size is larger than that of chatbots and algorithms in terms of perceived warmth (robots: $\bar{r} = -.382$; algorithms: $\bar{r} = -.169$, $p = .023$; chatbots: $\bar{r} = -.054$, $p < .001$) and perceived humanlikeness (robots: $\bar{r} = -.639$; algorithms: $\bar{r} = -.373$, $p = .015$; chatbots: $\bar{r} = -.147$, $p < .001$), with the exception that chatbots also outperform algorithms in terms of perceived humanlikeness ($p = .032$). Web Appendix K provides all p-values for these tests.

A column-wise comparison of the customer responses reveals some differences within customer perceptions and appraisals (Table 4, superscripts a, b, and c). For AAs overall, the negative effect size is smaller for perceived competence ($\bar{r} = -.135$), followed by perceived warmth ($\bar{r} = -.227$, $p = .014$) and then by perceived humanlikeness ($\bar{r} = -.380$, $p = .008$). Furthermore, the negative effect size is smaller for firm-related appraisals ($\bar{r} = -.074$) than for agent-related appraisals ($\bar{r} = -.193$, $p < .001$). These patterns apply to all AA types except for chatbots, where the three perceptions do not differ from each other. Web Appendix K provides all p-values for these tests.

Robot, Chatbot, Algorithm, and Overall Models

Table 5 shows the meta-regression results for robots, chatbots, algorithms, and AAs overall, explaining 51.1%, 33.9%, 33.1%, and 29.8% of the variance in the effect sizes, respectively. Web Appendix L provides the results for the control contingencies.

Robot model. The results support that name (b = .217, $p = .045$; H_{1a}), humanlike appearance (b = .031, $p = .009$; H_2), verbal-linguistic intelligence (b = .050, $p = .007$; H_{3a}), visual-spatial intelligence (b = .028, $p = .004$; H_5), and processing-speed intelligence (b = .069, $p = .004$; H_6) weaken the negative effect sizes of robots, whereas a social identity experience strengthens their negative effect sizes (b = −.389, $p < .001$; H_9). Unexpectedly, social intelligence (b = .030, $p = .077$; H_{4a}) and embarrassing experience (b = .061, $p = .460$; H_{8a}) have no effect.

Chatbot model. The results support that name (b = .189, $p < .001$; H_{1b}), verbal-linguistic intelligence (b = .091, $p < .001$; H_{3b}), and embarrassing experience (b = .401, $p = .007$; H_{8b}) weaken the negative effect sizes of chatbots. Unexpectedly, social intelligence has no effect (b = −.022, $p = .232$; H_{4b}). Although not hypothesized, humanlike appearance strengthens the negative effect sizes of chatbots (b = −.045, $p = .002$).

Algorithm model. The results support that utilitarian context weakens (b = .330, $p < .001$; H_{10}) and high-expertise role strengthens (b = −.149, $p = .008$; H_{11}) the negative effect sizes of algorithms. Unexpectedly, logic-mathematical intelligence has no effect (b = −.014, $p = .278$; H_7).

All AA type models. Regarding AA-type-independent contingencies, a negative outcome valence weakens the negative effect sizes of robots (b = .168, $p = .043$), chatbots (b = .192, $p = .006$), and algorithms (b = .155, $p < .001$). Time weakens the negative effect sizes of chatbots (b = .198, $p = .030$) and algorithms (b = .049, $p = .045$) but has no effect for robots (b = .095, $p = .335$).

Table 4. Pooled Effect Sizes Vary by AA Type and Customer Response.

AA Type	Robots			Chatbots			Algorithms			AAs Overall			Differences Across AA Types[1]
Customer Responses	No. of Effects k (Sample Size n)	Pooled Effect Size (SE)[2]	*p*	No. of Effects k (Sample Size n)	Pooled Effect Size (SE)[2]	*p*	No. of Effects k (Sample Size n)	Pooled Effect Size (SE)[2]	*p*	No. of Effects k (Sample Size n)	Pooled Effect Size (SE)[2]	*p*	RO = Robots CH = Chatbots AL = Algorithms
Perceptions													
Humanlikeness	15 (1,440)	−.639 (.099)[a]	<.001	17 (3,909)	−.147 (.070)	.051	28 (4,321)	−.373 (.080)[a]	<.001	60 (9,670)	−.380 (.065)[a]	<.001	RO < AL < CH
Warmth	23 (4,922)	−.382 (.097)[b]	<.001	18 (3,801)	−.054 (.043)	.227	25 (3,832)	−.169 (.049)[b]	.002	66 (12,555)	−.227 (.048)[b]	<.001	RO < AL = CH
Competence	21 (5,173)	−.199 (.094)[c]	.045	27 (4,320)	−.132 (.042)	.004	40 (9,090)	−.087 (.057)[b]	.133	88 (18,583)	−.135 (.040)[c]	<.001	—
Appraisals													
Agent-related	20 (4,893)	−.205 (.081)[a]	.018	38 (8,376)	−.237 (.048)[a]	<.001	96 (18,559)	−.152 (.047)[a]	<.001	154 (31,827)	−.193 (.032)[a]	<.001	—
Firm-related	83 (18,957)	−.057 (.043)[b]	.188	93 (27,373)	−.060 (.028)[b]	.037	96 (22,319)	−.087 (.020)[b]	<.001	272 (68,650)	−.074 (.018)[b]	<.001	—
Intentions													
Agent-related	21 (14,041)	−.180 (.058)	.005	N.A.[4]	N.A.[4]	N.A.[4]	21 (4,435)	−.245 (.049)[a]	<.001	42 (18,476)	−.217 (.038)	<.001	—
Firm-related	65 (23,984)	−.172 (.053)	.002	60 (16,359)	−.087 (.042)	.041	44 (7,778)	−.061 (.032)[b]	.060	169 (48,122)	−.120 (.028)	<.001	—
Behaviors													
Agent-related	3 (458)	−.084 (.260)	.777	N.A.[4]	N.A.[4]	N.A.[4]	48 (10,214)	−.028 (.048)	.570	51 (10,672)	−.021 (.048)	.670	—
Firm-related	11 (15,128)	−.077 (.043)	.105	7 (4,814)	−.059 (.111)	.612	6 (41,841)	−.160 (.107)	.190	24 (61,783)	−.014 (.052)	.788	—
Self-related	2 (219)	−.078 (.069)	.458	1 (70)	−.331 (na[4])	N.A.[4]	14 (1,329)	−.101 (.109)	.369	17 (1,618)	−.038 (.090)	.680	—
Responses Overall	**264 (89,215)**	**−.167 (.033)**	**<.001**	**261 (69,022)**	**−.122 (.029)**	**<.001**	**418 (123,718)**	**−.105 (.020)**	**<.001**	**943 (281,956)**	**−.127 (.015)**	<.001	**RO < (CH = AL)[3]**

[1]This column indicates differences ($p < .05$) in the pooled effect sizes across AA types (row-wise comparison). These differences were tested in the respective subsamples (e.g., studies examining perceived humanlikeness). We regressed the Fisher's z-transformed effect sizes on two dummies (representing robots, chatbots, and algorithms) and used the same model specification as for effect size pooling (i.e., weighted random-effects multilevel model). Web Appendix K provides all *p*-values for these tests.

[2]Differing superscripts a, b, and c indicate differences ($p < .05$) in the pooled effect sizes across customer responses within a response category, for example, customer perceptions (column-wise comparison). These differences were tested in the respective subsamples (e.g., studies examining the effect of robots vs. HAs on customer perceptions). Hereby, we proceeded analogously to the row-wise comparison with dummies representing the focal customer responses (e.g., perceived humanlikeness, warmth, competence). Web Appendix K provides all *p*-values for these tests.

[3]For the responses overall, the negative effect size is larger for robots ($\bar{r} = -.167$) than for algorithms ($\bar{r} = -.105$; $p = .044$), but not different from that for chatbots ($\bar{r} = -.122$; $p = .157$). Testing robots against the two digital AA types together shows that the negative effect size for robots ($\bar{r} = -.167$) is larger than for the other two AA types ($\bar{r} = -.110$; $p = .042$).

[4]The pooled effect size could not be estimated because the number of effect sizes k in this cell is 0 or 1.

Notes: The pooled effect sizes are based on random-effects multilevel models with the maximum likelihood (ML) estimator, using the inverse variance method with Fisher's z transformation of correlations.

Table 5. Meta-Regression Results with Unique Contingency Effects Across Robots, Chatbots, and Algorithms.

Contingencies[a] DV: Effect Sizes on Customer Responses	**Robots**			**Chatbots**			**Algorithms**			**AAs Overall**		
	Coefficient[c]	**SE**[c]	***p***	**Coefficient**[c]	**SE**[c]	***p***	**Coefficient**[c]	**SE**[c]	***p***	**Coefficient**[c]	**SE**[c]	***p***
Intercept	−.692	.337	.041	−.379	.350	.281	−.445	.252	.078	−.753	.147	<.001
Name	.217	.108	.045	.189	.051	<.001	.036	.055	.515	.114	.036	.002
Humanlike appearance	.031	.012	.009	−.045	.014	.002	.013	.017	.452	.009	.008	.242
Task-related intelligence												
Verbal-linguistic intelligence	.050	.018	.007	.091	.028	.001	[.005]	[.015]	[.745]	.020	.011	.081
Social intelligence	[.030]	[.017]	[.077]	−.022	.019	.232	−.005	.014	.726	.007	.011	.505
Visual-spatial intelligence	.028	.010	.004	[.033]	[.028]	[.246]	−.008	.016	.619	.020	.007	.004
Processing-speed intelligence	.069	.024	.004	−.014	.017	.385	.013	.016	.408	.036	.010	<.001
Logic-mathematical intelligence	−.009	.022	.681	.019	.033	.566	−.014	.013	.278	.006	.010	.525
Context												
Embarrassing experience	.061	.082	.460	.401	.147	.007	.008	.051	.876	.082	.042	.052
Social identity experience	−.389	.068	<.001	.051	.081	.528	.064	.053	.233	−.079	.038	.039
Utilitarian context	.085	.049	.085	.001	.059	.992	.330	.045	<.001	.176	.029	<.001
High-expertise role	[−.102]	[.092]	[.266]	−.016	.151	.915	−.149	.056	.009	−.107	.040	.007
Negative outcome valence	.168	.082	.043	.192	.069	.006	.155	.042	<.001	.158	.034	<.001
Time	.095	.098	.335	[.198]	[.091]	[.030]	[.049]	[.024]	[.045]	.038	.026	.147
Controls[d]												
Customer contingencies	Included			Included			Included			Included		
Method contingencies	Included			Included			Included					
R^2	51.1%			33.9%			33.1%			29.8%		
N (k)[e]	103 (264)			72 (261)			153 (418)			327 (943)		

[a]Customer responses (all models) and AA types (AA total model) were included as dummy variables but not presented for brevity. These are presented in Web Appendix L, Table W14.
[b]The humanlike appearance of chatbots and algorithms can be assessed only when they are visualized. Hence, we included visualization (1 = yes, 0 = no) as a control variable in the analyses for the chatbot (b = −.056, $p = .272$), algorithm (b = −.032, $p = .550$), and AAs overall (b = −.050, $p = .206$) models.
[c]The values for substantive contingencies with multicollinearity issues refer to ex post tests and are presented in square brackets.
[d]The control contingencies regarding customer and method were included. For the sake of brevity, they are presented in Web Appendix L.
[e]N and k refer to the number of studies and effects, respectively. One study reports separate effects for both robots and chatbots and thus, is considered in both subsamples. Hence, the summed number of studies (N = 328 = 103 + 72 + 153) is larger than the actual overall number depicted in the table (N = 327).
Notes: σ_b^2 = between-study variance (robots: .033, chatbots: .011, algorithms: .032, AAs overall: .033), σ_w^2 = within-study variance (robots: .034, chatbots: .035, algorithms: .026, AAs overall: .036). Unstandardized coefficients are shown. The contingency variable time is mean-centered.

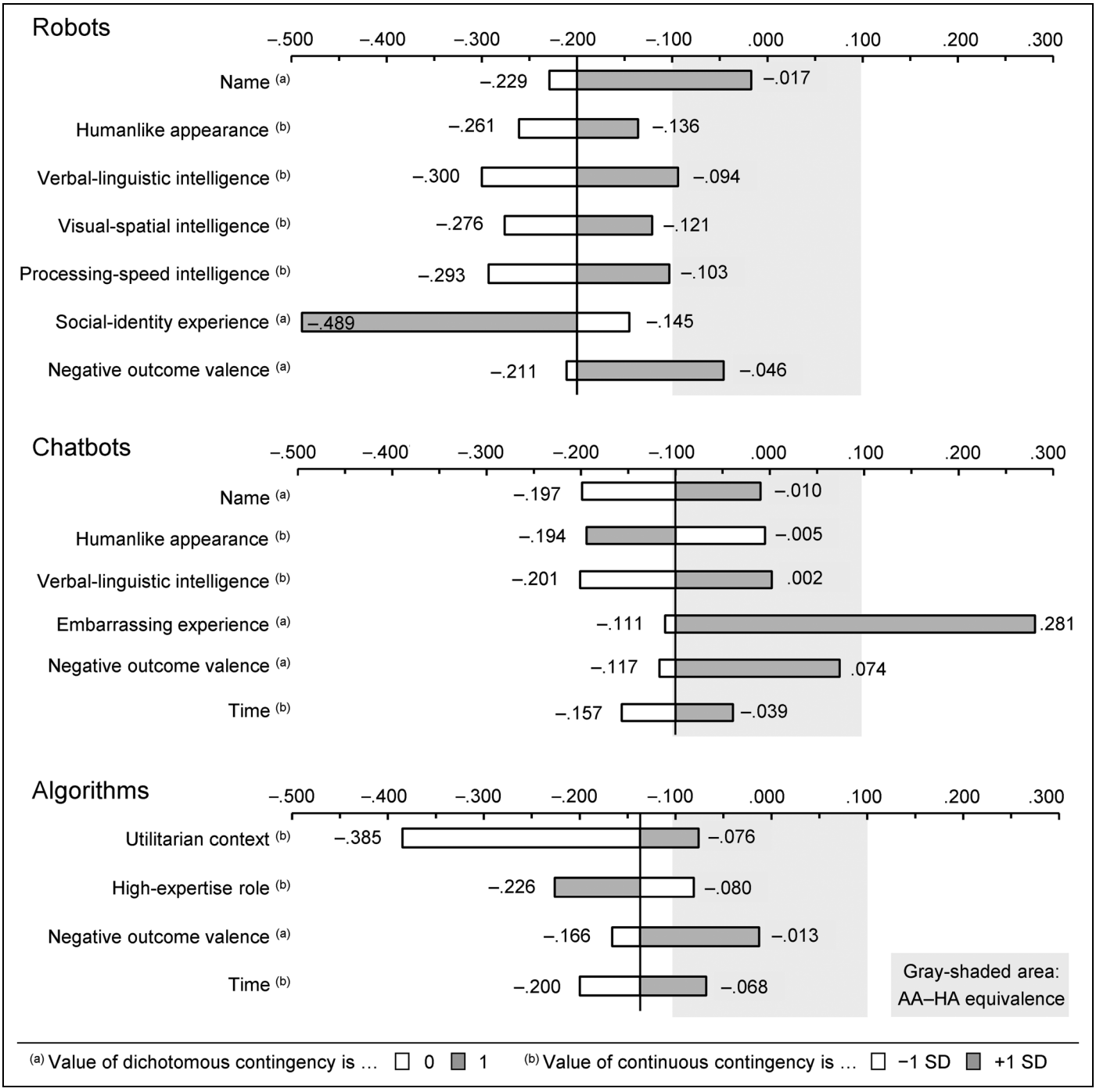


Figure 2. AAs Are Equivalent to HAs Under Several Conditions.

Notes: The values for each contingency are predicted when all nonfocal contingencies are at their means. The vertical axis is the mean effects size when all contingencies are at their means (robots: −.200; chatbots: −.100; algorithms: −.137). The predictions for time (chatbots and algorithms models) are based on ex post tests where the control contingencies were excluded due to multicollinearity issues. Reading example: The mean effect size of robots (vs. human agents) on customer responses is −.200 (vertical line). A name brings the effect size up to −.017. The lack of a name brings the effects size down to −.229.

Prediction of Effect Sizes

Using the estimated regression functions for the three AA types, we predict the effect sizes $\hat{r}$ for two levels of the relevant ($p < .05$) substantive contingencies (binary variables at their coded levels, continuous variables at ±1 SD of the variable's mean), freezing the nonfocal contingencies at their means (Melnyk, Carrillat, and Melnyk 2022). Figure 2 depicts the predicted values. Using Cohen's (1988) value of $\hat{r} = |.1|$ a threshold for a small effect, an AA is equivalent when $\hat{r} \geq -.1$ (i.e., when the AA is disfavored by .1 or less).

The results show that robots are equivalent to HAs if they have a name ($\hat{r} = -.017$), if the task requires verbal-linguistic intelligence ($\hat{r} = -.094$), and if the outcome valence is negative

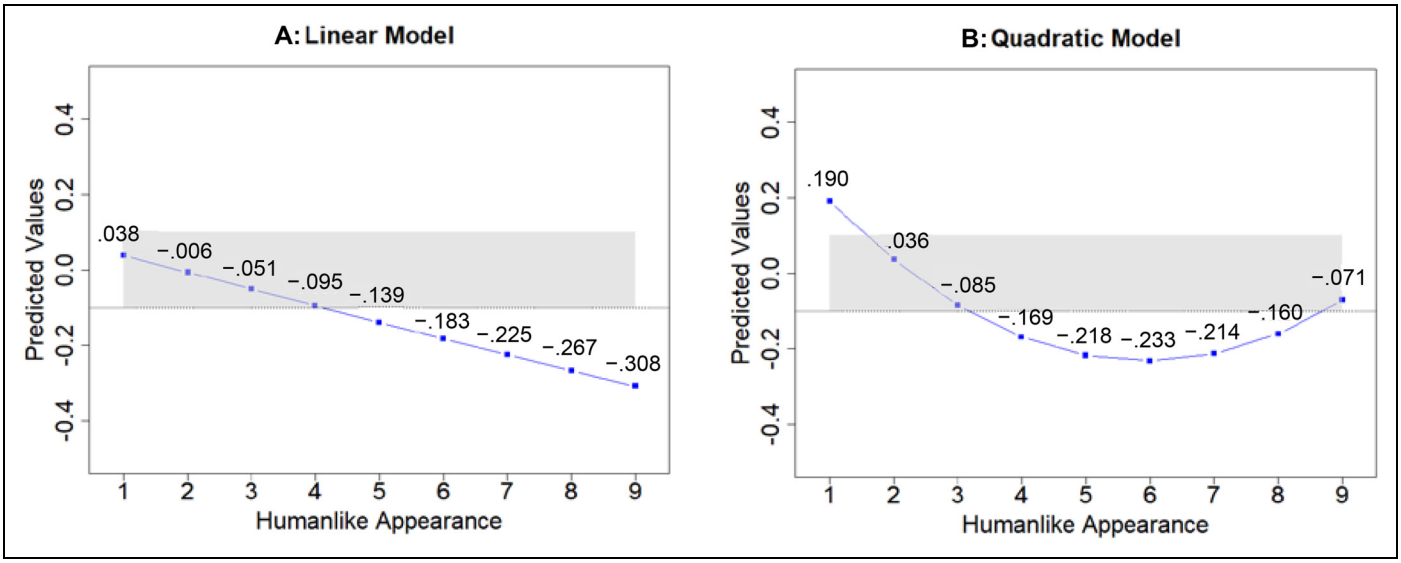


Figure 3. A U-Curve Qualifies the Negative Effect of a Humanlike Appearance for Chatbots.
Notes: The vertical axis is fixed at the mean effects size (r = −.100) when controlling for all contingency variables.

($\hat{r} = -.046$); they are essentially equivalent if a task requires processing-speed intelligence ($\hat{r} = -.103$). Chatbots are equivalent to HAs if they have a name ($\hat{r} = -.010$), if their humanlike appearance is low ($\hat{r} = -.005$), if the task requires verbal-linguistic intelligence ($\hat{r} = .002$), if the outcome valence is negative ($\hat{r} = .074$), and if they were used more recently ($\hat{r} = -.039$). In embarrassing contexts, chatbots are even superior to HAs ($\hat{r} = .281$). Algorithms are equivalent to HAs in utilitarian contexts ($\hat{r} = -.076$), in low-expertise roles ($\hat{r} = -.080$), if the outcome valence is negative ($\hat{r} = -.013$), and if they were used more recently ($\hat{r} = -.068$).

Nonlinear Effects of Humanlike Appearance

We explore the nonlinear effects of humanlike appearance by specifying a quadratic function for it in the models for the three AA types. The results for robots ($b = -.027$, $p = .598$; $c = .005$, $p = .241$) and algorithms ($b = -.013$, $p = .846$; $c = .003$, $p = .693$) show no effects, but they do for chatbots ($b = -.209$, $p < .001$; $c = .018$, $p < .001$), with the linear term indicated by b and the quadratic term by c. Figure 3 shows the predicted curves for chatbots. They indicate that the linear model (Panel A) is qualified by the quadratic model (Panel B), such that the overall negative slope shows a U-shaped curve. Web Appendix M provides further details (i.e., model-free evidence).

Discussion and Theoretical Implications

Consolidated knowledge about the human equivalence of the three AA types in marketing roles, as well as the contingencies of their human equivalence, is limited, leaving it unclear whether the findings apply equally to robots, chatbots, and algorithms (Table 1). We provide the first comprehensive assessment, offering several new insights.

Human Equivalence of Robots, Chatbots, and Algorithms

As a baseline, AAs are inferior to HAs ($\bar{r} = -.127$) in the overall view but are near the equivalence threshold. Comparing the AA types reveals that robots fall short of HAs the most (Table 4, row-wise comparison). These findings are in line with Zehnle, Hildebrand, and Valenzuela (2025), but our nuanced view of single customer responses reveals that robots' relative inferiority relates only to perceived humanlikeness and warmth. Presumably, due to their physical form of existence, it can be seen that robots are machines and therefore not humanlike or warm. Chatbots and algorithms equally outperform robots in the overall view. The relative parity of the digital AAs is surprising, as chatbots' conversational mode of communication mimics an online talk with a real human, which makes them appear more humanlike than algorithms. However, algorithms provide information on command and are well received when serving this purpose. A general algorithm aversion, often described in the literature (Dietvorst, Simmons, and Massey 2015), is thus not supported.

Comparing customer responses reveals an overarching pattern (Table 4, column-wise comparison): (1) AAs are most inferior for agent-related perceptions, (2) they are more inferior for agent- than for firm-related appraisals, and (3) they are equivalent to HAs for behaviors, be they agent-, firm-, or self-related. This pattern challenges the common assumption that AAs will remain inferior to human employees in the near future (Xiao and Kumar 2021). It also enriches prior meta-analyses, which do not distinguish between agent- and firm-related responses (Table 1). AAs are only inferior with respect to agent-related upstream responses, supporting the idea that customers harbor reservations about AAs. However, when it comes to assessing the firm and engaging in managerially relevant downstream responses (i.e., behaviors), customers value the output of AAs and choose or buy from them as if they were interacting with HAs.

Contingencies That Affect Equivalence

Our findings reveal idiosyncratic contingencies for the human equivalence of each AA type, with minor overlaps only (Table 5). They provide the following new insights.

Contingencies are AA-type-specific and not generalizable across AAs. This insight refutes the idea of generic or mixed-sample meta-analyses implying that "one size fits all." For example, the negative effect of a high-expertise role (Zehnle, Hildebrand, and Valenzuela 2025) applies only to algorithms, not to conversational AA types. We also qualify conflicting meta-analytic results (e.g., on the effect of a utilitarian vs. hedonic context; Blut and Wang 2020; Ma, Fan, and Mattila 2024), because they may veil AA-type-specific effects. The results also challenge two of our expectations about generalizability. First, an embarrassing context only benefits chatbots. The null effect for robots is surprising and contradicts single-study results (e.g., Holthöwer and Van Doorn 2023). One explanation is that robots are physical AAs acting in real space. Customers may feel less ashamed around them but will feel so while near human bystanders. The respective studies may have not aimed to account for such conceivable effects. Second, time only benefits chatbots and algorithms, not robots, which qualifies the positive effect of time in a mixed AA sample meta-analysis (Zehnle, Hildebrand, and Valenzuela 2025). Robots are difficult to design, and the period covered for them (from 2016 in our data) may be too short to see improvements. Hence, robots may still be somewhat unfamiliar to consumers (Van Doorn, Odekerken-Schröder, and Spohrer 2025).

The same contingency can have opposite effects in different AA types. A humanlike appearance increases the human equivalence of robots but shows a U-curve with a negative tendency for chatbots. Prior meta-analyses could not unveil such effects because many captured humanlikeness as perceived by customers rather than as an actual design feature (e.g., Blut et al. 2021; Blut, Wünderlich, and Brock 2024; Li et al. 2023). Even when testing a humanlike design, they mixed several facets (e.g., virtual face, character, name; Zehnle, Hildebrand, and Valenzuela 2025) or treated it as a categorical construct with equivocal results (Kilani and Rajaobelina 2024; Ma, Fan, and Mattila 2024). Examining nonlinear effects across AA types enriches the discussion of an uncanny valley (e.g., Mende et al. 2019), which does not seem to exist for robots and comes with a negative tendency and an earlier drop for chatbots. Chatbots' core purpose is to talk to customers in a virtual space. Visualizing them with an icon as a placeholder may best serve this purpose. A more humanlike look, especially a cartoonish or stylized one (i.e., a medium level), is distracting.

Generalizing AI is undue. This insight bears two facets. First, some prior meta-analyses generalize AAs as AI (e.g., Aguiar-Costa et al. 2022), although AAs may or may not be AI-enabled. We show that the AI required for a task is a contingency for AAs' equivalence with HAs, and this contingency may be more or less relevant. Second, the few meta-analyses that use AI as a contingency draw on a general, dichotomous concept of higher versus lower intelligence (Blut, Wünderlich, and Brock 2024; Ladeira, Perin, and Santini 2023). By considering Pantano and Scarpi's (2022) taxonomy of different AI types required for a task, we reveal that increasing AAs' general intelligence is not beneficial per se. Instead, different types of intelligence benefit different types of AAs. Robots can benefit the most because they engage in conversations (where increasing verbal-linguistic intelligence is helpful) and move in real space (where increasing visual-spatial and processing-speed intelligence is helpful).

Implications for Theory Building

The meta-regression results (Table 5) reveal two overarching patterns (Table 6). First, equivalence is affected by contingencies that can make AAs more humanlike. They represent verbal human cues (name and verbal-linguistic intelligence), which benefit robots and chatbots, and visual human cues (humanlike appearance), which benefit robots and harm chatbots. Second, equivalence is affected by contingencies that can make AAs' machine characteristics salient. These characteristics refer to the advantages and disadvantages of AAs over HAs (Table 2). Speed and reliability can benefit robots (in tasks related to visual-spatial or processing-speed intelligence) and algorithms (in utilitarian contexts). A lack of cognitive flexibility can benefit all AA types (for a negative outcome valence) but also harm algorithms (in a high-expertise role). A lack of agency can benefit chatbots (for an embarrassing experience). A lack of empathy can harm robots (for a social identity experience).

These patterns provide novel insights that enrich the ASP concept (Table 6). The core idea of this concept is that customers exposed to an AA feel the presence of a social entity (Van Doorn et al. 2017). This feeling can be triggered by verbal or nonverbal human cues (Short, Williams, and Christie 1976). Such cues cause customers to perceive AAs as human, despite their better knowledge (Holthöwer and Van Doorn 2023). This concept presumes that the humanlike nature of AAs overlays their machine nature. However, the two-partite configuration of contingencies suggests that customers may value both making AAs more humanlike and making their machine characteristics salient. Following this logic, we posit that customers perceive AAs as hybrid beings that bear a coexisting social presence (i.e., the feeling of interacting with a humanlike entity) and an automated presence (i.e., the feeling of interacting with a machine entity). The relevance of both dimensions depends on AA type.

Social presence only clearly benefits one AA type; for others, it can be irrelevant or even harmful (Table 6). This insight suggests that the merits of social presence assumed in ASP are not generalizable across AA types. It also puts into perspective meta-analytic findings, which suggest that social presence has positive effects on various customer responses (Blut et al. 2021; Blut, Wünderlich, and Brock 2024). The respective

Table 6. The Key Findings for the Contingencies Enrich the Concept of ASP.

Enriching ASP	Robots	Chatbots	Algorithms
AAs are hybrid beings with a social presence and an automated presence. Their roles depend on AAs' modes of communication and forms of existence.	Physical	Conversational	Digital / Informational
Social Presence	**Some contingencies make AAs more humanlike.**		
Social presence only matters for conversational AAs:			
• Verbal human cues benefit both of them.	• Name (+)	• Name (+)	
• Visual human cues benefit conversational AAs with a physical form but harm those with a digital form.	• Verbal-linguistic intelligence (+) • Humanlike appearance (+)	• Verbal-linguistic intelligence (+) • Humanlike appearance (−)[a]	
Automated Presence	**Other contingencies make AAs' machine characteristics salient.**		
Automated presence matters for all AA types, but different machine characteristics become salient:			
• AAs' advantages, their speed and reliability, benefits physical and informational AAs.	• Speed, reliability (+)		• Speed, reliability (+)
• AAs can even benefit from their disadvantages, their lack of cognitive flexibility (all AA types)[b] and lack of agency (conversational-digital AAs).	• Lack of cognitive flexibility (+)	• Lack of cognitive flexibility (+)	• Lack of cognitive flexibility (+, −[b])
• AAs' core disadvantage, their lack of empathy, can only harm physical AAs.	• Lack of empathy (−)	• Lack of agency (++)	

[a]This negative effect is qualified by a U-curve.
[b]Only for informational AAs, a lack of cognitive flexibility can also be harmful, namely when they take high-expertise roles.
Notes: The signs indicate the effect of the contingencies on an AA type's equivalence to HAs: "+" increases an AA type's human equivalence, "++" makes an AA type superior (threshold $r \gtrsim .1$), and "−" decreases an AA type's human equivalence.

meta-analyses lack a human benchmark, overlooking the fact that HAs can also benefit from a higher social presence, for example, through facial expressions (Short, Williams, and Christie 1976). Hence, when using HAs as a benchmark, the relevance of social presence looms smaller than currently assumed.

Automated presence is relevant for all AAs. Single AA types can benefit from AAs' advantageous machine characteristics. Disadvantageous machine characteristics are less harmful than presumed and can benefit all AAs or even make one AA type superior to HAs; only a lack of empathy remains problematic in one case (Table 6). These results enrich the ASP concept, which downplays the role of AAs' machine presence as a means to create a social presence (Van Doorn et al. 2017). Accordingly, recent meta-analyses have included AAs' social presence, but not their automated presence (e.g., Blut et al. 2021). Our results suggest that AAs' automated presence is an independent construct with multiple positive effects, depending on AA type and the task or context where it replaces a human employee.

In summary, our findings challenge the implicit assumption in the concept of ASP that AAs' social presence is key to success. Technological advancements may have led to an overrating of the need to make AAs more humanlike. This may not always be necessary, but AAs can be appropriate substitutes for HAs just because they are perceived as machines.

Managerial Implications

When should marketers consider AAs as an alternative to HAs? AAs alleviate labor shortages (Song et al. 2022) and increase efficiency (Xiao and Kumar 2021), calling for substitution when customer responses are equivalent (Figure 2). Our recommendations in Table 7 relate to design features and the tasks and contexts in which the use of AAs is desirable, as customers consider them equivalent ($\hat{r} \gtrsim -.1$). We add the tasks and contexts in which the use of AAs is acceptable, as efficiency gains may compensate for slightly lower customer receptivity ($-.2 \lesssim \hat{r} \lesssim -.1$). Since replacing human staff raises ethical concerns, firms should prioritize cases in which AAs relieve employees from physical or mental loads.

Robots

We recommend humanlike design features that incur low costs or strictly serve a given task. A useful low-cost feature is to give robots a name that helps customers engage them in conversations. Robots should also look humanlike, but this measure alone does not bring them on par with HAs and may be costly. Thus, designers should add low-cost humanlike elements (e.g., hair) and otherwise focus on elements necessary to fulfill a task. For example, wheels help robots move around, but fully functional legs are not necessary.

Table 7. The Use of the Three AA Types Is Recommended in Various Marketing Roles.

AA Type	Recommendations on Design Features	Recommendations on Tasks and Contexts: When the Use of AAs Is …	
		Desirable	Acceptable
Robots	• Give robots a name. • Give robots a humanlike look, but only for features that incur low costs or serve a purpose.	• Repetitive tasks that require processing-speed intelligence. Optimize robots' dexterity, but design simple interfaces. • Tasks that require verbal-linguistic intelligence. One option to fully exploit robots' potential in these tasks is to train them to ask customers for their preferences and store this information in the cloud to provide a customized service. For teaching and training, ensure privacy to avoid public denigration.	Tasks that require visual-spatial intelligence.
Chatbots	• Give chatbots a name. • When visualizing chatbots, use a simple icon. Never use a cartoonish or stylized avatar.	• Tasks that require verbal-linguistic intelligence. The most appropriate application is a one-on-one learning context, but many sales-related tasks also fall into this category (e.g., providing recommendations). One way to fully exploit chatbots' potential in these tasks is to ask customers for their preferences and store this information in the cloud. • Embarrassing context. Place chatbots prominently in the app or on the website and tout their anonymity benefits.	Virtually all conditions. Hence, the ubiquitous use of chatbots is recommended.
Algorithms	It is not necessary to give algorithms a name or a humanlike look.	• Utilitarian context. When risk is involved, offer human contact as an additional option. For utilitarian products with a high touch, offer chatbots in addition to algorithms that allow chatting for experiential pleasure. • Low-expertise role. Focus on increasing ease of use through simple interfaces and tracking users' purchase histories and locations.	In many conditions, except for hedonic contexts and high-expertise roles. Hence, a wide use of algorithms is recommended.
All AA types	—	• Sending bad news to customers. When firms seek to provide recovery, we recommend robots (for tangible interactions) or chatbots (for digital interactions). They can provide an explanation and a solution but forward customers to an HA in highly emotional situations. Algorithms should not be used for recovery.	In all acceptable conditions, only use AAs when efficiency gains compensate for lower customer receptivity.

The use of robots is desirable when a task requires processing-speed intelligence or verbal-linguistic intelligence. Processing-speed intelligence is key for repetitive jobs that need to be performed rapidly, where robots can relieve HAs from physical loads. Examples include room service in hotels, taking standardized orders, preparing coffee, or clearing dishes. To ensure speed in repetitive jobs, developers should optimize robots' dexterity, for example, always holding the cup at the correct angle when frothing milk. Long conversations are neither required nor expected for these tasks. Hence, we recommend simple interfaces, limited to basic conversations related to robots' duties. For room service, this could be a touchscreen in the room to place a predetermined set of orders and another touchscreen on the robot to send the robot out. For tasks requiring verbal-linguistic intelligence, one option to fully exploit robots' potential is to train them to ask customers about their preferences and combine this information with knowledge retrieved from the cloud. This could help customize the conversation. One example is an intelligent robotic waiter (Wang et al. 2022), which not only knows the nutritional details of the menu (Collins 2020) but also remembers a regular guest's intolerance and proactively recommends a particular dish. The tasks that require the highest verbal skills are teaching and training, such as giving a lecture (Li et al. 2016) or playing exergames with elderly people (Čaić et al. 2020). However, since customers may fear public denigration, we recommend robotic instructors only in private settings (e.g., at home or in single booths).

The use of robots is acceptable for tasks that require visual-spatial intelligence, such as cleaning or ironing at home (Kim, Schmitt, and Thalmann 2019) and guiding travelers at airports (Hwang et al. 2022). However, firms should assess whether efficiency gains compensate for lower customer acceptance and whether robots can relieve employees' physical or mental load, which may be greater for cleaning than for ironing.

Chatbots

We recommend only one humanlike design feature that incurs low costs. Firms should give chatbots a name to facilitate conversations. Furthermore, when visualizing chatbots, they should use a simple icon, as in the case of ChatGPT. A very humanlike avatar is a second-best option and requires more design resources. In any case, firms should avoid stylized avatars with cartoon-like human features, such as Microsoft's Clippy.

The use of chatbots is desirable when a task requires verbal-linguistic intelligence or in embarrassing contexts. The highest verbal fluency is required for one-on-one learning contexts, such as improving English grammar skills or providing online lectures. However, many sales-related tasks also fall into this category (e.g., discussing recommendations for leisure time activities, the pros and cons of a product, personal food preferences). As with robots, one way to fully exploit chatbots' potential is to train them to ask customers for their preferences, store them in the cloud, and retrieve them when necessary (e.g., remember color preferences for apparel). When customers buy embarrassing products, such as diarrhea pills or incontinence diapers, chatbots outperform HAs. Hence, firms selling embarrassing products should place chatbots prominently on their apps or websites, presenting them as a way to ensure anonymity.

The use of chatbots is acceptable even when none of the desirable conditions apply. This is because chatbots never fall below the $-.2$ threshold for $\hat{r}$ (Figure 2). These results call for ubiquitous use whenever the expected efficiency gains of up to 30% (Xiao and Kumar 2021) compensate for lower acceptance, specifically when HAs are relieved of mental stress (e.g., handling multiple inquiries at the same time).

Algorithms

Algorithms do not need humanlike design features. Their use is desirable in utilitarian contexts or low-expertise roles. Algorithms lend themselves to utilitarian contexts, such as calculating a distance or forecasting wait times. Nevertheless, even in these contexts, emotions can be involved. This applies to decisions that entail risk, such as protecting one's home (insurance), deciding on a retirement plan (banking), and buying baby food (grocery shopping). In these contexts, HAs should be in place as an option, such as a live chat or a hotline, to convey warmth and reduce fear. Some people also seek experiential pleasure even when buying utilitarian products (e.g., high-tech/high-touch products such as laptops). Hence, an algorithm should provide basic information (e.g., on prices and features), but customers should have the option to connect with a chatbot for exploration purposes (e.g., "Explore playful features of our laptops"). This is because chatbots that follow an entertaining script can provide experiential pleasure (Sands et al. 2021).

Examples of algorithms in low-expertise roles include determining the right clothing sizes in online shops, selling tickets, and giving advice on how to protect oneself from bad weather. Since these tasks involve standard inquiries and simple requests, firms should focus on increasing the ease of use of algorithms for customers. An interface that takes commands easily and provides simple outputs, for example, through a one-click repeat order or an autofill function, can serve this purpose. Ease of use can also be increased by tracking users' purchase history (e.g., to suggest an appropriate outfit and clothing size) and location (e.g., to provide location-based advice on how to be protected from bad weather).

The use of algorithms is generally acceptable, except for hedonic contexts or high-expertise roles, where they fall below the $-.2$ threshold for $\hat{r}$. Except for these cases, we recommend the wide use of algorithms if efficiency gains compensate for lower customer acceptance, specifically when HAs are relieved of mental stress (e.g., repetitive inquiries).

All AA Types

The use of all AAs is desirable for communicating bad news to customers. It prevents HAs from mental stress (e.g., when algorithms or chatbots deny a loan application or a club membership) and even physical harm from aggressive customers (e.g., when robots hand out parking tickets or deny guests without tickets access to a festival). In these examples, firms do not seek to recover customers, and AAs' work is done.

In other cases—for example, when informing customers about problems (i.e., a service failure)—firms should recover customers. These efforts require verbal skills, particularly when exploring solutions for customers. Firms can exploit the verbal talents of robots (for tangible interactions) and chatbots (for digital interactions) to explain situations and offer solutions. For example, robots could clarify that access to theme park rides is controlled based on crowding and that, for safety reasons, ten guests are admitted every five minutes. Chatbots can explain why a check-in failed (Zhang et al. 2023), for example, by informing a guest that their code expired and that a new one is being sent. Nevertheless, negative outcomes differ in severity (McQuilken 2010), and customers vary in their levels of anger during incidents (Gelbrich 2010). Hence, robots or chatbots should be trained to recognize highly emotional situations (Huang and Rust 2024) and forward customers to an HA to prevent escalation. As algorithms lack verbal skills, we do not recommend them for service recovery.

Research Agenda

This meta-analysis is a starting point for answering further pressing questions regarding AA–HA comparisons. Our

Table 8. Research Agenda with Pressing Questions and Recommended Examinations.

Question	Recommended Examinations
1. Are AAs preferred when the human benchmark is less attractive?	• Compare AAs to a situation when HAs are unavailable (e.g., a restaurant or bank branch that has closed) or when HA availability is limited (e.g., food delivery during the daytime only, inquiries during the week only, a longer wait time on a hotline with HAs). • Compare discounted AAs to full-price HAs: – Use different discount levels, between 0% (the same price as HAs) and 100% (for free), to determine how many customers would switch to an AA. – Determine the optimal discount level for AAs (i.e., where the ratio of savings for HAs to costs for AAs is highest).
2. How can the social presence of robots and chatbots be further increased, and when is this beneficial?	• Test whether and when novel personality traits are beneficial for AAs: – Test the fit between personality and role. Examples are "servile" robots as domestic helpers, "pushy" chatbots as personal trainers, or "compassionate" chatbots as buddies. – Explore how to create specific personalities (e.g., by using phrases like "Certainly," "Sure," "Yes, madam," or "Always at your service" for servile AAs). • Test whether and when cues of human imperfection are beneficial for AAs: – Test verbal cues for robots and chatbots, such as typing as slowly as humans, or at least slower than machines (vs. typing fast), and using (vs. not using) colloquial language or mild curse words. – Test visual cues for robots, such as making them look older (vs. younger). – Test boundary conditions for the cues, e.g., a high-expertise (vs. low-expertise) role, a utilitarian (vs. hedonic) context, or a public (vs. private) setting.
3. Under which further conditions do AAs benefit from their automated presence?	• Test further conditions when AAs can leverage their speed and reliability. Examples are customers with (vs. without) time pressure or in a business (vs. private) context. • Test further conditions when AAs can leverage their lack of cognitive flexibility (i.e., follow a standardized procedure that neglects customers' uniqueness). This may be the case when customers have low (vs. high) levels of status or low (vs. high) cognitive resources. • Test further conditions when conversational AAs can leverage their lack of agency (i.e., they cannot judge). This may be the case when customers could harm HAs during the interaction, for example, when using robots as sparring partners or when filing a vindictive complaint to chatbots. When interacting in real space (i.e., with robots), control for the effects of bystanders, who may still judge customers. • Clarify whether a lack of bystanders is a precondition for robots to benefit from their lack of agency in embarrassing contexts.
4. What is the interplay of social presence and automated presence?	• Test when social presence is more important than (or completely overrides) automated presence (e.g., in online therapy, for personal buddies) and vice versa (e.g., for performance tracking). • Test the interaction effects of the cues related to AAs' social presence and the conditions related to AAs' automated presence. For example, slowing down chatbot typing (i.e., a cue of social presence) may be appreciated only when customers are not under time pressure (i.e., when an automated presence is less relevant).
5. How receptive are customers to AAs across different countries?	• Test the acceptance of AAs (vs. HAs) across cultures: – Test if uncertainty avoidance generally hinders AAs' acceptance due to the associated risks or fosters acceptance when AAs make helpful forecasts. – Test the interaction effects of cultural dimensions with variables related to social and automated presence. For example, power distance may foster the acceptance of robots with a "servile" personality, and collectivism may foster the acceptance of "compassionate" chatbots. • Test the effects of economic factors, such as national wealth and labor shortages, on the acceptance of AAs and make predictions about future usage (e.g., national wealth driving AA acceptance).

research agenda comprises five key questions, based on this current study's findings, its limitations, and managerial considerations (Table 8). First, this study used comparable HAs as a control condition for AAs. Future research should examine whether AAs are preferred when this human benchmark is less attractive, either by making HAs unavailable or by discounting AAs. Second, our results suggest that increasing social presence plays a role in conversational AAs. Recent research shows beneficial effects of social presence. Specifically, robots with a conscientious (vs. extraverted) personality enhance customer engagement in a banking (vs. restaurant) context (Balaji et al. 2024). In contrast, chatbots that use conversational fillers—signs of humans' imperfection (Lee and Yi 2025)—are shown to decrease purchase intentions when used by for-profit (vs. nonprofit) firms (Liu, Liu, and Zhu 2024). Future research should examine novel human cues (e.g., other personality traits, imperfection) to increase AAs' social presence and test when they are beneficial. Third, our results suggest that automated presence can play a role in each AA type. Future research could identify further conditions under which AAs benefit from their machine characteristics. Fourth, we suggest that AAs are hybrid beings with a social and an automated presence. Future research could further contribute to theory

building by examining the interplay between the two dimensions. The fifth recommendation relates to examining country differences in the receptiveness of AAs (vs. HAs). Only 60% of the included studies reported participants' nationalities. Half of these studies were conducted in North America, limiting cultural variety. Research has shown that cultural differences affect consumers' reactions to AAs (Pitardi et al. 2023); thus, the acceptance of AAs across cultural dimensions and other country-level factors requires further study.

Acknowledgments

The authors would like to thank Chiara Orsingher for providing valuable input early in the research process. The authors would also like to thank the *JM* review team for their fruitful feedback and for helping develop this manuscript throughout the revision process.

Coeditor

Detelina Marinova

Associate Editor

Dhruv Grewal

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Katja Gelbrich https://orcid.org/0000-0003-0147-3064
Holger Roschk https://orcid.org/0000-0003-2521-8367
Sandra Miederer https://orcid.org/0009-0008-5550-3373
Alina Kerath https://orcid.org/0009-0008-1282-9592

References

Aguiar-Costa, Laura M., Carlos A.X.C. Cunha, Wallysson K.M. Silva, and Nelsio R. Abreu (2022), "Customer Satisfaction in Service Delivery with Artificial Intelligence: A Meta-Analytic Study," *Revista de Administração Mackenzie*, 23 (6), http://dx.doi.org/10.1590/1678- 6971/eramd220003.en.

Alexander, Ralph A., Michael J. Scozzaro, and Lawrence J. Borodkin (1989), "Statistical and Empirical Examination of the Chi-Square Test for Homogeneity of Correlations in Meta-Analysis," *Psychological Bulletin*, 106 (2), 329–31.

Allard, Thomas and Katherine White (2015), "Cross-Domain Effects of Guilt on Desire for Self-Improvement Products," *Journal of Consumer Research*, 42 (3), 401–19.

Appel, Markus, David Izydorczyk, Silvana Weber, Martina Mara, and Tanja Lischetzke (2020), "The Uncanny of Mind in a Machine: Humanoid Robots as Tools, Agents, and Experiencers," *Computers in Human Behavior*, 102, 274–86.

Balaji, M.S., Priyanka Sharma, Yangyang Jiang, Xiya Zhang, Steven T. Walsh, Abhishek Behl, and Kokil Jain (2024), "A Contingency-Based Approach to Service Robot Design: Role of Robot Capabilities and Personalities," *Technological Forecasting and Social Change*, 201, 123257.

Beattie, Austin, Autumn P. Edwards, and Chad Edwards (2020), "A Bot and a Smile: Interpersonal Impressions of Chatbots and Humans Using Emoji in Computer-Mediated Communication," *Communication Studies*, 71 (3), 409–27.

Belanche, Daniel, Luis V. Casaló, Jeroen Schepers, and Carlos Flavián (2021), "Examining the Effects of Robots' Physical Appearance, Warmth, and Competence in Frontline Services: The Humanness-Value-Loyalty Model," *Psychology & Marketing*, 38 (12), 2357–76.

Blut, Markus, Arezou Ghiassaleh, and Cheng Wang (2023), "Testing the Performance of Online Recommendation Agents: A Meta-Analysis," *Journal of Retailing*, 99 (3), 440–59.

Blut, Markus and Cheng Wang (2020), "Technology Readiness: A Meta-Analysis of Conceptualizations of the Construct and its Impact on Technology Usage," *Journal of the Academy of Marketing Science*, 48 (4), 649–69.

Blut, Markus, Cheng Wang, Nancy V. Wünderlich, and Christian Brock (2021), "Understanding Anthropomorphism in Service Provision: A Meta-Analysis of Physical Robots, Chatbots, and Other AI," *Journal of the Academy of Marketing Science*, 49 (4), 632–58.

Blut, Markus, Nancy V. Wünderlich, and Christian Brock (2024), "Facilitating Retail Customers' Use of AI-Based Virtual Assistants: A Meta-Analysis," *Journal of Retailing*, 100 (2), 293–315.

Čaić, Martina, João Avelino, Dominik Mahr, Gaby Oderkerken-Schröder, and Alexandre Bernardino (2020), "Robotic Versus Human Coaches for Active Aging: An Automated Social Presence Perspective," *International Journal of Social Robotics*, 12 (4), 867–82.

Castelo, Noah, Johannes Boegershausen, Christian Hildebrand, and Alexander P. Henkel (2023), "Understanding and Improving Consumer Reactions to Service Bots," *Journal of Consumer Research*, 50 (4), 848–63.

Choi, Miju, Youngjoon Choi, Seongseop Kim, and Frank Badu-Baiden (2023), "Human vs Robot Baristas During the COVID-19 Pandemic: Effects of Masks and Vaccines on Perceived Safety and Visit Intention," *International Journal of Contemporary Hospitality Management*, 35 (2), 469–91.

Choi, Youngjoon, Miju Choi, Munhyang Oh, and Seongseop Kim (2020), "Service Robots in Hotels: Understanding the Service Quality Perceptions of Human-Robot Interactions," *Journal of Hospitality Marketing & Management*, 29 (6), 613–35.

Clegg, Melanie, Reto Hofstetter, Emanuel de Bellis, and Bernd H. Schmitt (2023), "Unveiling the Mind of the Machine," *Journal of Consumer Research*, 51 (2), 342–61.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.

Collins, Galen R. (2020), "Improving Human–Robot Interactions in Hospitality Settings," *International Hospitality Review*, 34 (1), 61–79.

Cooper, Harris (2017), *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*, Applied Social Research Methods Series, 5th ed. Sage.

Crolic, Cammy, Felipe Thomaz, Rhonda Hadi, and Andrew T. Stephen (2022), "Blame the Bot: Anthropomorphism and Anger in

Customer–Chatbot Interactions," *Journal of Marketing*, 86 (1), 132–48.

Dahl, Darren W., Rajesh V. Manchanda, and Jennifer J. Argo (2001), "Embarrassment in Consumer Purchase: The Roles of Social Presence and Purchase Familiarity," *Journal of Consumer Research*, 28 (3), 473–81.

De Keyser, Arne and Werner H. Kunz (2022), "Living and Working with Service Robots: A TCCM Analysis and Considerations for Future Research," *Journal of Service Management*, 33 (2), 165–96.

Desideri, Lorenzo, Cristina Ottavani, Massimiliano Malavsi, and Roberto Di Marzi (2019), "Emotional Processes in Human-Robot Interaction During Brief Cognitive Testing," *Computers in Human Behavior*, 90, 331–42.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey (2015), "Algorithm Aversion People Erroneously Avoid Algorithms After Seeing Them Err," *Journal of Experimental Psychology*, 144 (1), 114–26.

Dootson, Paula, Dominique A. Greer, Kate Letheren, and Kate L. Daunt (2023), "Reducing Deviant Consumer Behaviour with Service Robot Guardians," *Journal of Services Marketing*, 37 (3), 276–86.

Drouin, Michelle, Susan Sprecher, Robert Nicola, and Taylor Perkins (2022), "Is Chatting with a Sophisticated Chatbot as Good as Chatting Online or FTF with a Stranger?" *Computers in Human Behavior*, 128, 107100.

El Halabi, Malak and Olivier Trendel (2024), "Just Name it: The Act of Naming Humanoid Service Robots Decreases Perceived Eeriness and Increases Repurchase Intent," *Journal of Service Research*, 28 (1), 131–49.

Epley, Nicholas, Adam Waytz, and John T. Cacioppo (2007), "On Seeing Human: A Three-Factor Theory of Anthropomorphism," *Psychological Review*, 114 (4), 864–86.

Fiske, Susan T., Amy J.C. Cuddy, Peter and, and Glick (2007), "Universal Dimensions of Social Cognition: Warmth and Competence," *Trends in Cognitive Sciences*, 11 (2), 77–83

Flavián, Carlos, Russell W. Belk, Daniel Belanche, and Luis V. Casaló (2024), "Automated Social Presence in AI: Avoiding Consumer Psychological Tensions to Improve Service Value," *Journal of Business Research*, 175, 114545.

Frank, Darius-Aurel and Tobias Otterbring (2023), "Being Seen… by Human or Machine? Acknowledgment Effects on Customer Responses Differ Between Human and Robotic Service Workers," *Technological Forecasting and Social Change*, 189, 122345.

Garvey, Aaron M., TaeWoo Kim, and Adam Duhachek (2022), "Bad News? Send an AI. Good News? Send a Human," *Journal of Marketing*, 87 (1), 1–16.

Gelbrich, Katja (2010), "Anger, Frustration, and Helplessness After Service Failure: Coping Strategies and Effective Informational Support," *Journal of the Academy of Marketing Science*, 38 (5), 567–85.

Go, Eun and Shyan S. Sundar (2019), "Humanizing Chatbots: The Effects of Visual, Identity and Conversational Cues on Humanness Perceptions," *Computers in Human Behavior*, 97, 304–16.

Golder, Peter N., Marnik G. Dekimpe, Jake T. An, Harald J. van Heerde, Darren S. Kim, and Joseph W. Alba (2023), "Learning from Data: An Empirics-First Approach to Relevant Knowledge Generation," *Journal of Marketing*, 87 (3), 319–36.

Gopinath, Krishnan and Dharun Kasilingam (2023), "Antecedents of Intention to Use Chatbots in Service Encounters: A Meta-Analytic Review," *International Journal of Consumer Studies*, 47 (6), 2367–95.

Harrer, Mathias, Pim Cuijpers, Toshi A. Furukawa, and David D. Ebert (2021), *Doing Meta-Analysis with R: A Hands-on Guide*. CRC Press.

Higgins, Julian P.T. and Simon G. Thompson (2002), "Quantifying Heterogeneity in a Meta-Analysis," *Statistics in Medicine*, 21 (11), 1539–58.

Holthöwer, Jana and Jenny van Doorn (2023), "Robots Do Not Judge: Service Robots Can Alleviate Embarrassment in Service Encounters," *Journal of the Academy of Marketing Science*, 51, 767–84.

Hoyer, Wayne D., Deborah J. MacInnis, and Rik Pieters (2024), *Consumer Behavior*, 8th ed. Cengage Learning.

Huang, Ming-Hui and Roland T. Rust (2018), "Artificial Intelligence in Service," *Journal of Service Research*, 21 (2), 155–72.

Huang, Ming-Hui and Roland T. Rust (2021), "Engaged to a Robot? The Role of AI in Service," *Journal of Service Research*, 24 (1), 30–41.

Huang, Ming-Hui and Roland T. Rust (2024), "The Caring Machine: Feeling AI for Customer Care," *Journal of Marketing*, 88 (5), 1–23.

Huang, Guanxiong and Sai Wang (2023), "Is Artificial Intelligence More Persuasive Than Humans? A Meta-Analysis," *Journal of Communication*, 73 (6), 552–62.

Hunter, John E. and Frank L. Schmidt (2004), *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, 2nd ed. Sage.

Hwang, Jinsoo, Heather M. Kim, Kyu-Hyeon Joo, and Won S. Lee (2022), "How to Form Rapport with Information Providers in the Airport Industry: Service Robots Versus Human Staff," *Asia Pacific Journal of Tourism Research*, 27 (8), 891–906.

Ivanov, Stanislav, Craig Webster, and Faruk Seyitoğlu (2023), "Humans and/or Robots? Tourists' Preferences Towards the Humans–Robots Mix in the Service Delivery System," *Service Business*, 17 (1), 195–231.

Jin, Jianna, Jesse Walker, and Rebecca W. Reczek (2024), "Avoiding Embarrassment Online: Response to and Inferences About Chatbots When Purchases Activate Self-Presentation Concerns," *Journal of Consumer Psychology*, 35 (2), 185–202.

Joshi, Naveen (2022), "3 Key Differences Between AI And Robotics," *Forbes* (January 16), https://www.forbes.com/sites/naveenjoshi/2022/01/16/3-key-differences-between-ai-and-robotics/.

Kätsyri, Jari, Klaus Förger, Meeri Mäkäräinen, and Tapio Takala (2015), "A Review of Empirical Evidence on Different Uncanny Valley Hypotheses: Support for Perceptual Mismatch as One Road to the Valley of Eeriness," *Frontiers in Psychology*, 6, 390.

Khan, Uzma, Ravi Dhar, and Klaus Wertenbroch (2005), "A Behavioral Decision Theory Perspective on Hedonic and Utilitarian Choice," in *Inside Consumption: Consumer Motives, Goals, and Desires*, S. Ratneshwar and David G. Mick, eds. Routledge, 166–87.

Kilani, Nour and Lova Rajaobelina (2024), "Impact of Live Chat Service Quality on Behavioral Intentions and Relationship Quality: A Meta-Analysis," *International Journal of Human-Computer Interaction*, 40 (7), 1558–85.

Kim, Boyoung, Ewart J. de Visser, and Elizabeth Phillips (2022), "Two Uncanny Valleys: Re-Evaluating the Uncanny Valley Across the Full Spectrum of Real-World Human-Like Robots," *Computers in Human Behavior*, 135, 107340.

Kim, Seo Y., Bernd H. Schmitt, and Nadia M. Thalmann (2019), "Eliza in the Uncanny Valley: Anthropomorphizing Consumer Robots Increases Their Perceived Warmth but Decreases Liking," *Marketing Letters*, 30 (1), 1–12.

Kuen, Leonie, Daniel Westmattelmann, Maike Bruckes, and Gerhard Schewe (2023), "Who Earns Trust in Online Environments? A Meta-Analysis of Trust in Technology and Trust in Provider for Technology Acceptance," *Electron Markets*, 33 (1).

Ladeira, Wagner, Marcelo G. Perin, and Fernando Santini (2023), "Acceptance of Service Robots: A Meta-Analysis in the Hospitality and Tourism Industry," *Journal of Hospitality Marketing & Management*, 32 (6), 694–716.

Larkin, Connor, Caitlin Drummond Otten, and Joseph Árvai (2022), "Paging Dr. JARVIS! Will People Accept Advice from Artificial Intelligence for Consequential Risk Management Decisions?" *Journal of Risk Research*, 25 (4), 407–22.

Lee, Heekyung and Youjae Yi (2025), "Humans vs. Service Robots as Social Actors in Persuasion Settings," *Journal of Service Research*, 28 (1), 150–67.

Leng, Minmin, Peng Liu, Ping Zhang, Mingyue Hu, Haiyan Zhou, Guichen Li, Huiru Yin, and Li Chen (2019), "Pet Robot Intervention for People with Dementia: A Systematic Review and Meta-Analysis of Randomized Controlled Trials," *Psychiatry Research*, 271, 516–25.

Li, Bin, Yanhong Chen, Luning Liu, and Bowen Zheng (2023), "Users' Intention to Adopt Artificial Intelligence-Based Chatbot: A Meta-Analysis," *Service Industries Journal*, 43 (15–16), 1117–39.

Li, Jamy, René Kizilcec, Jeremy N. Bailenson, and Wendy Ju (2016), "Social Robots and Virtual Agents as Lecturers for Video Instruction," *Computers in Human Behavior*, 55 (Part B), 1222–30.

Liadeli, Georgia, Francesca Sotgiu, and Peeter W. Verlegh (2023), "A Meta-Analysis of the Effects of Brands' Owned Social Media on Social Media Engagement and Sales," *Journal of Marketing*, 87 (3), 406–27.

Lipsey, Mark W. and David B. Wilson (2001), *Practical Meta-Analysis*. Sage.

Liu, Guilin, Maggie W. Liu, and Qichao Zhu (2024), "Hmm, the Effect of AI Conversational Fillers on Consumer Purchase Intentions," *Marketing Letters* (published online November 23), https://doi.org/10.1007/s11002-024-09760-4 .

Longoni, Chiara, Andrea Bonezzi, and Carey K. Morewedge (2019), "Resistance to Medical Artificial Intelligence," *Journal of Consumer Research*, 46 (4), 629–50.

Longoni, Chiara and Luca Cian (2022), "Artificial Intelligence in Utilitarian vs. Hedonic Contexts: The 'Word-of-Machine' Effect," *Journal of Marketing*, 86 (1), 91–108.

Luo, Xueming, Siliang Tong, Zheng Fang, and Zhe Qu (2019), "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases," *Marketing Science*, 38 (6), 937–47.

Ma, Chang, Alei Fan, and Anna S. Mattila (2024), "Decoding the Shared Pathways of Consumer Technology Experience in Hospitality and Tourism: A Meta-Analysis," *International Journal of Hospitality Management*, 118, 103685.

McLean, Graeme, Kofi Osei-Frimpong, and Jennifer Barhorst (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, 312–28.

McQuilken, Lisa (2010), "The Influence of Failure Severity and Employee Effort on Service Recovery in a Service Guarantee Context," *Australasian Marketing Journal*, 18 (4), 214–21.

Mehta, Pooja, Charles Jebarajakirthy, Haroon I. Maseeh, Anubha Anubha, Raiswa Saha, and Komal Dhanda (2022), "Artificial Intelligence in Marketing: A Meta-Analytic Review," *Psychology & Marketing*, 39 (11), 2013–38.

Melnyk, Vladimir, François A. Carrillat, and Valentyna Melnyk (2022), "The Influence of Social Norms on Consumer Behavior: A Meta-Analysis," *Journal of Marketing*, 86 (3), 98–120.

Mende, Martin, Maura L. Scott, Jenny van Doorn, Dhruv Grewal, and Ilana Shanks (2019), "Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses," *Journal of Marketing Research*, 56 (4), 535–56.

Miller, Elizabeth J., Yong Z. Foo, Paige Mewton, and Amy Dawel (2023), "How Do People Respond to Computer-Generated Versus Human Faces? A Systematic Review and Meta-Analyses," *Computers in Human Behavior Reports*, 10, 100283.

Mori, Masahiro, Karl F. MacDorman, and Norri Kageki (2012), "The Uncanny Valley [From the Field]," *IEEE Robotics & Automation Magazine*, 19 (2), 98–100.

Önkal, Dilek, Paul Goodwin, Mary Thomson, Sinan Gönül, and Andrew Pollock (2009), "The Relative Influence of Advice from Human Experts and Statistical Methods on Forecast Adjustments," *Journal of Behavioral Decision Making*, 22 (4), 390–409.

Pan, Yadong, Haruka Okada, Toshiaki Uchiyama, and Kenji Suzuki (2015), "On the Reaction to Robot's Speech in a Hotel Public Space," *International Journal of Social Robotics*, 7 (5), 911–20.

Pantano, Eleonora and Daniele Scarpi (2022), "I, Robot, You, Consumer: Measuring Artificial Intelligence Types and Their Effect on Consumers Emotions in Service," *Journal of Service Research*, 25 (4), 583–600.

Pitardi, Valentina, Boris Bartikowski, Victoria-Sophie Osburg, and Vignesh Yoganathan (2023), "Effects of Gender Congruity in Human-Robot Service Interactions: The Moderating Role of Masculinity," *International Journal of Information Management*, 70, 102489.

Pitardi, Valentina, Jochen Wirtz, Stefanie Paluch, and Werner H. Kunz (2022), "Service Robots, Agency, and Embarrassing Service Encounters," *Journal of Service Management*, 33 (2), 389–414.

Prentice, Catherine and Mai Nguyen (2020), "Engaging and Retaining Customers with AI and Employee Service," *Journal of Retailing and Consumer Services*, 56, 102186.

Roschk, Holger and Katja Gelbrich (2017), "Compensation Revisited: A Social Resource Theory Perspective on Offering a Monetary Resource After a Service Failure," *Journal of Service Research*, 20 (4), 393–408.

Ruan, Yanya and József Mezei (2022), "When Do AI Chatbots Lead to Higher Customer Satisfaction Than Human Frontline Employees in Online Shopping Assistance? Considering Product Attribute Type," *Journal of Retailing and Consumer Services*, 68, 103059.

Ruiz-Equihua, Daniel, Jaime Romero, Luis V. Casaló, and Sandra M. C. Loureiro (2023), "Smart Speakers and Customer Experience in Service Contexts," *Psychology & Marketing*, 40 (11), 2326–40.

Ryoo, Yuhosua, Yongwoog A. Jeon, and WooJin Kim (2024), "The Blame Shift: Robot Service Failures Hold Service Firms More Accountable," *Journal of Business Research*, 171, 114360.

Saha, Sudip (2023), "Autonomous Agents Market Outlook (2023–2033)," (accessed January 5, 2024), https://www.futuremarketinsights.com/reports/autonomous-agents-market.

Sanders, Tracy, Alexandra D. Kaplan, Ryan Koch, Michael Schwartz, and Peter A. Hancock (2019), "The Relationship Between Trust and Use Choice in Human-Robot Interaction," *Human Factors*, 61 (4), 614–26.

Sands, Sean, Carla Ferraro, Colin L. Campbell, and Hsiu-Yuan Tsao (2021), "Managing the Human–Chatbot Divide: How Service Scripts Influence Service Experience," *Journal of Service Management*, 32 (2), 246–64.

Schmitt, Bernd H. (1999), "Experiential Marketing," *Journal of Marketing Management*, 15 (1–3), 53–67.

Schoeffer, Jakob, Yvette Machowski, and Niklas Kuehl (2021), "Perceptions of Fairness and Trustworthiness Based on Explanations in Human vs. Automated Decision-Making," arXiv, https://doi.org/10.48550/arXiv.2109.05792 .

Short, John, Ederyn Williams, and Bruce Christie (1976), *The Social Psychology of Telecommunications*. John Wiley & Sons.

Song, Hanqun, Yao-Chin Wang, Huijun Yang, and Emily Ma (2022), "Robotic Employees vs. Human Employees: Customers' Perceived Authenticity at Casual Dining Restaurants," *International Journal of Hospitality Management*, 106, 103301.

Srinivasan, Raji and Gülen Sarial-Abi (2021), "When Algorithms Fail: Consumers' Responses to Brand Harm Crises Caused by Algorithm Errors," *Journal of Marketing*, 85 (5), 74–91.

Stern, Steven E. and John W. Mullennix (2004), "Sex Differences in Persuadability of Human and Computer-Synthesized Speech: Meta-Analysis of Seven Studies," *Psychological Reports*, 94 (3 Pt 2), 1283–92.

Thomaz, Felipe, Carolina Salge, Elena Karahanna, and John Hulland (2020), "Learning from the Dark Web: Leveraging Conversational Agents in the Era of Hyper-Privacy to Enhance Marketing," *Journal of the Academy of Marketing Science*, 48 (1), 43–63.

Van Doorn, Jenny, Martin Mende, Stephanie M. Noble, John Hulland, Amy L. Ostrom, Dhruv Grewal, and J.A. Petersen (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.

Van Doorn, Jenny, Gaby Odekerken-Schröder, and Jim Spohrer (2025), "Robots Are Here to Stay: Time to Invest in a Future We Actually Want to Live In," *Journal of Service Research*, 28 (1), 3–8.

Veerbeek, Janne M., Anneli C. Langbroek-Amersfoort, Erwin E.H. van Wegen, Carel G.M. Meskers, and Gert Kwakkel (2017), "Effects of Robot-Assisted Therapy for the Upper Limb After Stroke," *Neurorehabilitation and Neural Repair*, 31 (2), 107–21.

Viechtbauer, Wolfgang (2010), "Conducting Meta-Analyses in R with the Metafor Package," *Journal of Statistical Software*, 36 (3), 1–48.

Wang, Sai and Guanxiong Huang (2024), "The Impact of Machine Authorship on News Audience Perceptions: A Meta-Analysis of Experimental Studies," *Communication Research*.

Wang, Yawei, Qi Kang, Shoujiang Zhou, Yuanyuan Dong, and Junqi Liu (2022), "The Impact of Service Robots in Retail: Exploring the Effect of Novelty Priming on Consumer Behavior," *Journal of Retailing and Consumer Services*, 68, 103002.

Wang, Cuicui, Yiyang Li, Weizhong Fu, and Jia Jin (2023), "Whether to Trust Chatbots: Applying the Event-Related Approach to Understand Consumers' Emotional Experiences in Interactions with Chatbots in E-Commerce," *Journal of Retailing and Consumer Services*, 73, 103325.

Wien, Anders H. and Alessandro M. Peluso (2021), "Influence of Human Versus AI Recommenders: The Roles of Product Type and Cognitive Processes," *Journal of Business Research*, 137, 13–27.

Wieseke, Jan, Anja Geigenmüller, and Florian Kraus (2012), "On the Role of Empathy in Customer-Employee Interactions," *Journal of Service Research*, 15 (3), 316–31.

Wirtz, Jochen, Paul G. Patterson, Werner H. Kunz, Thorsten Gruber, Vinh N. Lu, Stefanie Paluch, and Antje Martins (2018), "Brave New World: Service Robots in the Frontline," *Journal of Service Management*, 29 (5), 907–31.

Xiao, Li and V. Kumar (2021), "Robotics for Customer Service: A Useful Complement or an Ultimate Substitute?" *Journal of Service Research*, 24 (1), 9–29.

Xie, Zhaohan, Yining Yu, Jing Zhang, and Mingliang Chen (2022), "The Searching Artificial Intelligence: Consumers Show Less Aversion to Algorithm-Recommended Search Product," *Psychology & Marketing*, 39 (10), 1902–19.

Yalcin, Gizem, Sarah Lim, Stefano Puntoni, and Stijn M. van Osselaer (2022), "Thumbs Up or Down: Consumer Reactions to Decisions by Algorithms Versus Humans," *Journal of Marketing Research*, 59 (4), 696–717.

Yu, Shubin, Ji Xiong, and Hao Shen (2022), "The Rise of Chatbots: The Effect of Using Chatbot Agents on Consumers' Responses to Request Rejection," *Journal of Consumer Psychology* (34), 35–48.

Zehnle, Meike, Christian Hildebrand, and Ana Valenzuela (2025), "Not all AI Is Created Equal: A Meta-Analysis Revealing Drivers of AI Resistance Across Markets, Methods, and Time," *International Journal of Research in Marketing* (published online February 13), https://doi.org/10.1016/j.ijresmar.2025.02.005 .

Zhang, Lixuan, Iryna Pentina, and Yuhong Fan (2021), "Who Do You Choose? Comparing Perceptions of Human vs Robo-Advisor in the Context of Financial Services," *Journal of Services Marketing*, 35 (5), 634–46.

Zhang, Jiemin, Yimin Zhu, Jifei Wu, and Grace F. Yu-Buck (2023), "A Natural Apology is Sincere: Understanding Chatbots' Performance in Symbolic Recovery," *International Journal of Hospitality Management*, 108, 103387.

Zhou, Qi, Bin Li, Lei Han, and Min Jou (2023), "Talking to a Bot or a Wall? How Chatbots vs. Human Agents Affect Anticipated Communication Quality," *Computers in Human Behavior*, 143, 107674.