The Impact of Anthropomorphism on Perceived Warmth and Competence of Artificial Intelligence and Willingness for Human-AI Cooperation

Authors: Cui Xunxuan, Qiao Ziteng, Liu Ning, Liu Ning

Date: 2026-02-04T11:49:41+00:00

Abstract

With the rapid development of artificial intelligence technology, human–AI collaboration has attracted increasing attention. Although anthropomorphism is widely regarded as an important strategy to promote human–AI interaction, its effects are unstable and can sometimes even be counterproductive. Through six progressive studies, this paper examines the impact of anthropomorphism on willingness to cooperate with AI, reveals its psychological mechanisms from a social-cognitive perspective, and further investigates the moderating role of perceived threat (including realistic threat and uniqueness threat). Studies 1a and 1b, using passive textual manipulation, found that anthropomorphism can indirectly enhance cooperation willingness by increasing perceived warmth. Study 2, using active imagination manipulation, found that anthropomorphism can simultaneously enhance perceived warmth-competence and cooperation willingness, and that perceived warmth-competence serves as the mediating mechanism through which anthropomorphism promotes the willingness for human–AI cooperation. Study 3a replicated the path whereby anthropomorphism influences cooperation willingness through perceived warmth-competence, and Study 3b further found that even under high realistic threat conditions, the mediating role of perceived warmth and competence still holds, with realistic threat exerting no moderating effect. Study 4 manipulated anthropomorphism through robot appearance images and found that appearance-based anthropomorphism negatively affects the willingness for human–AI cooperation, and this effect is moderated by perceived uniqueness threat. These findings help answer the key questions of why and when to anthropomorphize in human–AI cooperation, and provide theoretical implications for optimizing anthropomorphic design in artificial intelligence and promoting human–AI collaboration.

Full Text

The Impact of Anthropomorphism on Perceived Warmth-Competence of AI and Human-AI Cooperation Intention

CUI Xunxuan[1],[2], QIAO Ziteng[3], LIU Ning[1],[2]

[1] Faculty of Psychology, Shandong Normal University, Jinan 250358, China
[2] Shandong Provincial Key Laboratory of Brain Science and Mental Health, Jinan 250014, China
[3] School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China

Abstract

As artificial intelligence (AI) technology advances rapidly, human-AI collaborative relationships have attracted increasing attention. Although anthropomorphism is widely regarded as a key strategy to promote human-AI interaction, its effects remain unstable and sometimes counterproductive. This paper investigates the influence of anthropomorphism on human-AI cooperation intention through six sequential studies. From a social cognition perspective, we reveal its underlying psychological mechanisms and examine the moderating role of perceived threat (including realistic threat and uniqueness threat). Studies 1a and 1b, using passive textual manipulation, found that anthropomorphism indirectly promotes cooperation intention by enhancing perceived warmth. Study 2, employing active imagination manipulation, demonstrated that anthropomorphism simultaneously increases perceived warmth and competence, with both dimensions serving as mediators between anthropomorphism and cooperation intention. Study 3a replicated the mediation pathway through perceived warmth-competence, while Study 3b further revealed that this pathway remains robust even under high realistic threat conditions, with no moderating effect observed. Study 4 used robotic appearance images to manipulate anthropomorphism and found that appearance-based anthropomorphism negatively affects human-AI cooperation intention, moderated by perceived uniqueness threat. These findings address critical questions about *how* and *when* to anthropomorphize AI in collaborative contexts, providing theoretical guidance for optimizing AI anthropomorphic design to facilitate human-AI cooperation.

Keywords: anthropomorphism, human-AI collaboration, perceived warmth-competence, perceived threat, social cognition
Classification: B849: C91

1 Introduction

The rapid development of Artificial Intelligence (AI) is profoundly reshaping the cognitive landscape of human society. As AI evolves from narrow applications (e.g., AlphaGo, DeepSeek) toward Artificial General Intelligence (AGI;

Mitchell, 2024), the relationship between humans and AI has shifted from traditional instrumental interactions to more collaborative cognitive partnerships—a consensus reached in the international academic community (Du et al., 2024; NASEM, 2021; O' Neill et al., 2022; Xu et al., 2024). Consequently, exploring how humans and AI can coexist harmoniously represents a critical research agenda in the intelligent era.

A core dilemma in current human-AI interaction is that while technological capabilities enable complex task collaboration, public distrust and interaction resistance severely constrain the translation of technological potential into practical efficacy (Shariff et al., 2017; Yeomans et al., 2019). As AI systems increasingly exhibit human-like cognitive abilities, humanity faces both philosophical questions about "what it means to be human" and practical challenges about "how AI can become a partner." Addressing these challenges requires robust theoretical and empirical support. To resolve this impasse, AI anthropomorphism has emerged as a key research direction, with numerous studies positioning it as a core strategy to enhance human-AI interaction (e.g., Bartneck et al., 2009; Glikson & Woolley, 2020; Pelau et al., 2021; Qi et al., 2025; Roesler et al., 2021; Waytz et al., 2014; Xie et al., 2023; Zhou et al., 2023). Nevertheless, research also documents negative psychological consequences of anthropomorphism (Crolic et al., 2022; Ferrari et al., 2016; Mori, 1970; Vimalkumar et al., 2021; Yam et al., 2022). Therefore, investigating the mechanisms (*how* to anthropomorphize) and boundary conditions (*when* to anthropomorphize) is essential. This study examines the psychological mechanisms and boundaries of AI anthropomorphism from a social cognition perspective, aiming to facilitate human-AI interaction in the intelligent era.

1.1 Anthropomorphism in Human-AI Interaction

As AI' s emotional, behavioral, and cognitive capabilities increasingly approximate human abilities, its integration into society raises critical questions about risk, trust, and ethics. How humans effectively interact with AI has become a prominent topic across psychology, marketing, and computer science (Han et al., 2023; Schanke et al., 2021). To promote human-AI interaction, researchers have highlighted the importance of anthropomorphism, including its manifestation in appearance (Alabed et al., 2022; Ferrari et al., 2016), language (Xie et al., 2023; Zhou et al., 2023), movement or behavior (Eyssel et al., 2010; Qi et al., 2025; Yam et al., 2022; Zhou et al., 2023), and contextual framing (Onnasch & Roesler, 2020). In this research, AI anthropomorphism refers to the attribution of human characteristics, motivations, intentions, or mental states to non-human AI entities (Epley et al., 2007). This conceptualization encompasses not just isolated aspects of appearance, behavior, or personality, but rather the process of endowing non-human objects with human uniqueness (Epley et al., 2007; Xu et al., 2017).

Based on these characteristics, anthropomorphism research in human-AI interaction can be categorized into four types: appearance, communication, movement,

and context (Onnasch & Roesler, 2020). Studies show that AI anthropomorphism can yield both positive effects (Bartneck et al., 2009; Pelau et al., 2021; Qi et al., 2025; Xie et al., 2023; Zhou et al., 2023) and negative consequences (Crolic et al., 2022; Ferrari et al., 2016; Mori, 1970; Vimalkumar et al., 2021; Yam et al., 2022). Moreover, whether anthropomorphism proves beneficial depends on the method of anthropomorphization (Xie et al., 2023), interaction context (Onnasch & Roesler, 2020; Zhang et al., 2025), and user characteristics such as mindset (Han et al., 2023; Schanke et al., 2021), expectations (Khadpe et al., 2020), perceptions (Yam et al., 2021), and emotions (Crolic et al., 2022).

Overall, the benefits of AI anthropomorphism appear to outweigh its drawbacks (Roesler et al., 2021; Xie et al., 2023). Anthropomorphism enables AI to transcend simple instrumentality, becoming socially intelligent agents that fulfill human expectations (Roesler et al., 2021; Yu & Xu, 2020). Undeniably, anthropomorphism represents an effective human-centered approach to promoting human-AI interaction, making it crucial to clarify how anthropomorphism influences such interactions.

The human-AI relationship is undergoing a significant transformation from tool to partner (McKee et al., 2023; O' Neill et al., 2022; Xu et al., 2024), compelling us to consider how to cooperate with AI partners (O' Neill et al., 2022; Xu et al., 2024). Human-AI collaboration refers to the process where humans and AI work together to accomplish shared goals (O' Neill et al., 2022). Existing research supports that anthropomorphism facilitates human-AI interaction by making AI more likely to be attributed with human-like traits and mental states (Gray & Wegner, 2012), increasing trust (Glikson & Woolley, 2020; Waytz et al., 2014), acceptance (Harris-Watson et al., 2023), usage intention (Pelau et al., 2021), and interaction satisfaction (Xie et al., 2023).

In summary, whether anthropomorphism promotes human-AI collaboration and which types of anthropomorphism are effective remain unclear. Based on the overall positive effects of anthropomorphism on human-AI interaction, we propose **Hypothesis 1**: Anthropomorphism promotes human-AI cooperation intention.

1.2 Anthropomorphism and Warmth-Competence Social Cognition

Anthropomorphism' s effectiveness in promoting human-AI interaction stems from its ability to trigger a shift in human social cognition frameworks, transforming AI from an instrumental object into a human-like social cognition target. Building on this logic, this study examines anthropomorphism' s impact on human-AI collaboration from a social cognition perspective. Among various social cognition theories, the Stereotype Content Model (SCM) provides a classic and robust framework, proposing that warmth and competence constitute the two fundamental dimensions of social cognition (Fiske et al., 2002). Warmth relates to perceived intentions (e.g., friendliness, sincerity, morality), while competence concerns the ability to fulfill those intentions (e.g., intelligence, ambition,

efficiency) (Zuo et al., 2015). Although warmth and competence have typically been studied in relation to humans, researchers have found that non-human entities—including animals (Sevillano & Fiske, 2016), consumer brands (Kervyn et al., 2012), organizations (Aaker et al., 2010), and virtual agents (Bergmann et al., 2012)—are also subject to warmth and competence perceptions.

As human perceptions of AI shift from tool to partner, AI becomes a target of social cognition. Research demonstrates that the warmth-competence model applies to human perceptions and judgments of AI (e.g., Gilad et al., 2021; Khadpe et al., 2020; McKee et al., 2023; Reeves et al., 2020; Wu et al., 2025). For instance, humans typically perceive AI as high in competence but low in warmth (McKee et al., 2023; Wu et al., 2025), with these perceptions varying according to AI' s physical attributes, human-like appearance, and functional roles (Lee et al., 2011; McKee et al., 2023; Reeves et al., 2020). However, no study has systematically examined how anthropomorphism influences warmth and competence perceptions of AI.

Group categorization—or more precisely, in-group versus out-group perception —significantly influences warmth-competence perceptions (Fiske et al., 2002). Research shows that compared to out-groups, people evaluate in-group members as higher in both perceived warmth and competence (e.g., Fiske et al., 2002; Tian & Zhang, 2007). Eyssel and Kuchenbrandt (2012) found that human in-group preferences extend to robots: through manipulations of names and origins, robots can be categorized as in-group or out-group members, eliciting in-group favorability and stronger anthropomorphic tendencies toward in-group robots.

We propose that anthropomorphism is a key factor facilitating the reclassification of AI from out-group to in-group. The essence of anthropomorphism lies in perceiving human-like features in non-human agents (Epley et al., 2008), which helps humans view AI as in-group members (Haslam, 2006; Leyens et al., 2000), thereby increasing perceived warmth and competence. Supporting evidence shows that human-like traits influence warmth and competence perceptions of AI (Bergmann et al., 2012; Qi et al., 2025; Yam et al., 2021; Zhou et al., 2023). Consequently, we propose **Hypothesis 2a**: Anthropomorphism enhances perceived warmth and competence of AI.

Intergroup interaction research indicates that perceived warmth and competence are crucial determinants of social interaction intentions (Cuddy et al., 2011; Fiske et al., 2002). Perceived warmth relates to social acceptance and connection necessary for survival, while perceived competence involves status and ability judgments required for goal achievement (Fiske et al., 2007). Thus, AI' s perceived warmth and competence are critical for rapid decisions about collaboration. In human-AI collaboration contexts, researchers have found that AI systems can be perceived as warm and competent, and that these perceptions influence cooperative behavior (McKee et al., 2023). Furthermore, higher perceived warmth and competence predict greater cooperation intention (Khadpe et al., 2020; Piçarra & Giger, 2018). Related research has examined mediating mechanisms such as psychological closeness (Xie et al., 2023) and perceived em-

pathy (Pelau et al., 2021), with studies showing that warmth and competence mediate the effects of anthropomorphism on attitudes and continued usage intention (Qi et al., 2025; Zhou et al., 2023). Wu et al. (2025) similarly found that warmth and competence predict liking and usage intention toward large language model AI. This study investigates the role of perceived warmth and competence in the relationship between anthropomorphism and human-AI collaboration from a social cognition perspective. We hypothesize that AI anthropomorphism promotes cooperation intention by enhancing perceived warmth and competence. Therefore, we propose **Hypothesis 2b**: Perceived warmth and competence mediate the relationship between anthropomorphism and cooperation intention.

1.3 The Moderating Role of Perceived Threat

As discussed, AI anthropomorphism may foster in-group perception, enhance warmth-competence social cognition, and thereby promote human-AI collaboration. However, does this positive effect have boundary conditions? We propose that perceived threat from AI can moderate the influence of anthropomorphism on cooperation intention. AI has permeated all domains of human production and life, reshaping not only human-computer interaction patterns but also subtly influencing social cognition regarding self-identity and intergroup relations (Xu et al., 2024; Yogeeswaran et al., 2016). The potential threats posed by AI to human society remain a focal concern for both academia and the public.

Intergroup Threat Theory (Stephan et al., 2016) posits that out-groups can trigger realistic threat and uniqueness threat (also called symbolic threat). Realistic threat refers to potential damage by out-groups to in-groups' actual resources and survival interests, encompassing political power, economic benefits, social status, and personal safety (Zhang et al., 2009). Uniqueness threat concerns threats to in-group identity, cultural values, and group distinctiveness. In human-AI contexts, realistic threat manifests as AI replacing human jobs, posing public safety risks (e.g., algorithmic decision errors), and competing for scarce resources (McClure, 2018). Uniqueness threat is defined here as anthropomorphic AI blurring the boundary between human and machine during social interaction, challenging human self-perceptions of uniqueness (e.g., autonomous consciousness, emotional experience, moral judgment), thereby reducing evaluation of and interaction willingness with AI (Ferrari et al., 2016; Yogeeswaran et al., 2016).

Perceived threat is a key factor that accentuates and strengthens in-group/out-group distinctions (Stephan et al., 2002; Stephan et al., 2009). While anthropomorphism may promote AI in-group categorization and enhance cooperation intention through the warmth-competence pathway, intergroup threat can disrupt this process. When individuals perceive realistic threat from AI, AI is viewed as a direct competitor for resources, leading individuals to derogate out-groups and undermine human-AI cooperation to maintain in-group advantage and self-esteem (Golec de Zavala et al., 2019). Similarly, uniqueness threat

leads individuals to interpret AI' s anthropomorphic features as excessively approximating human essence, blurring human-machine boundaries (Yogeeswaran et al., 2016). This challenges human group distinctiveness and identity boundaries (Abbink & Harris, 2019; Ferrari et al., 2016). To defend group distinctiveness and maintain existing social identity (Brewer, 2007), individuals derogate and reject out-group AI (Riek et al., 2006), inhibiting human-AI cooperation. Therefore, we propose **Hypothesis 3**: Perceived threat moderates the effect of anthropomorphism on human-AI cooperation intention. Specifically, under low-threat conditions, anthropomorphism promotes cooperation intention; under high-threat conditions, anthropomorphism fails to promote or even reduces cooperation intention.

1.4 Research Overview

In summary, this study integrates the Warmth-Competence Social Cognition Model and Intergroup Threat Theory to investigate how AI anthropomorphism affects perceived warmth-competence and cooperation intention, testing whether perceived warmth-competence mediates this relationship, and examining the moderating role of perceived threat to identify boundary conditions. The core hypothesis is:

Anthropomorphism enhances perceived warmth and competence of AI, which in turn promotes human-AI cooperation, while perceived threat moderates this effect.

Specifically, six sequential studies test these hypotheses. Anthropomorphism manipulation methods include passive textual manipulation (Studies 1a, 1b), active imagination manipulation (Studies 2, 3a, 3b), and appearance manipulation (Study 4). Study 1 examines the basic effects of anthropomorphism (anthropomorphic vs. non-anthropomorphic) on perceived warmth-competence and cooperation intention. Study 2 tests the mediating mechanism through which anthropomorphism (high vs. low) influences cooperation intention via perceived warmth-competence. Study 3a replicates the mediation pathway, while Study 3b further explores the moderating role of realistic threat. Study 4 uses ecologically valid appearance manipulation to examine anthropomorphism' s effects and the moderating role of uniqueness threat.

2 Study 1: Passive Textual Manipulation of Anthropomorphism

Study 1 aimed to manipulate anthropomorphism through textual descriptions and examine its effects on AI' s perceived warmth-competence and human-AI cooperation intention.

2.1.1 Participants

We used G*Power 3.1 (Faul et al., 2009) to estimate the required sample size. For an independent samples* t*-test with a medium effect size of* $d^* = 0.5$, significance level $\alpha = 0.05$, and 90% statistical power, at least 172 participants were needed. To account for data exclusion, and following previous studies using similar manipulation paradigms (Zhao et al., 2024), we recruited participants through the credamo platform, with real-time exclusion of those failing attention checks or completing the survey in less than 120 seconds, and used rolling recruitment. Ultimately, 278 participants (179 female) voluntarily participated, with a mean age of 31.20 years ($SD = 8.80$). All participants provided informed consent after reading the experimental instructions and received compensation upon completion.

2.1.2 Design and Procedure

Study 1a employed a single-factor between-subjects design with two levels: anthropomorphic manipulation group versus non-anthropomorphic manipulation group (hereafter referred to as anthropomorphic and non-anthropomorphic groups). Participants were randomly assigned to one of the two groups and first read manipulation materials containing either anthropomorphic or non-anthropomorphic descriptions. The collaboration scenario was adapted from Harris-Watson et al. (2023), requiring participants to recall a reference team they had previously worked in or were currently part of. Participants then answered questions about their reference team, including whether they had worked in it, its size, and its type. After this brief survey, participants were asked to imagine that an AI teammate would join their team to handle performance evaluation tasks.

Anthropomorphic group AI material: "Hi there! My name is Archie, and I' m your intelligent HR assistant. I will be responsible for conducting this year' s employee performance evaluation at Chuhé Group. I will assess each employee' s performance data. Additionally, I will evaluate employees' five-minute self-statement videos. After reviewing the performance data and videos, I will make the final decision on this year' s employee performance evaluation independently and with full authority. This decision may affect employees' salaries, bonuses, promotion eligibility, and in some cases, even termination."

Non-anthropomorphic group AI material: "HRA is a new HR intelligent algorithm. This year' s employee performance evaluation at Chuhé Group will be conducted by algorithm HRA. Algorithm HRA will assess each employee' s performance data. Additionally, algorithm HRA will evaluate employees' five-minute self-statement videos. After reviewing the performance data and videos, the HR intelligent algorithm HRA will make the final decision on this year' s employee performance evaluation independently and with full authority. This decision may affect employees' salaries, bonuses, promotion eligibility, and in some cases, even termination."

This manipulation method followed established paradigms for textual anthropomorphism (Zhao et al., 2024). The anthropomorphic group featured an AI assistant with a name and first-person descriptions, while the non-anthropomorphic group used an objective, third-person description of an algorithm. All other AI descriptions were identical. Previous research indicates that direct presentation of anthropomorphic language can effectively induce anthropomorphism (Tam et al., 2013; Zhao et al., 2024). Following the manipulation, participants completed an anthropomorphism check ("To what extent does the HR intelligent assistant 'Archie' /intelligent algorithm 'HRA' remind you of human characteristics?") on a 7-point Likert scale (1 = not at all, 7 = very much).

After the manipulation check, both groups reported their cooperation intention with the AI using four items from the psychological acceptance dimension of the Human-AI Teams (HATS) newcomer acceptance scale (Harris-Watson et al., 2023) (e.g., "This AI could be a good teammate") on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Cronbach' s α for this scale was 0.92.

Participants then evaluated the AI' s perceived warmth and competence using Hu et al.' s (2021) AI perceived warmth-competence scale, which includes five items each for warmth (e.g., "This AI is friendly") and competence (e.g., "This AI is intelligent") on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Cronbach' s α was 0.93 for the warmth dimension and 0.79 for the competence dimension.

Finally, participants reported demographic information including gender and age. Given established links between gender and cooperation intention (Balliet et al., 2011), gender was included as a control variable in all analyses (hereafter, this will not be repeated).

2.1.3 Results and Analysis

Manipulation Check and Control Variables: An independent samples t-test on the anthropomorphism check showed that the anthropomorphic group ($M = 5.22$, $SD = 1.12$) scored significantly higher than the non-anthropomorphic group ($M = 4.90$, $SD = 1.28$), $t(276) = 2.25$, $p = 0.03$, Cohen' s $d = 0.27$, confirming the effectiveness of the manipulation. Age differences between groups were non-significant, $t(276) = 0.77$, $p = 0.44$. Gender was not significantly correlated with perceived warmth, competence, or cooperation intention, so it was excluded from subsequent analyses (Bernerth & Aguinis, 2016).

Effect of Anthropomorphism on Perceived Warmth-Competence: A 2 (anthropomorphic/non-anthropomorphic) × 2 (warmth/competence) mixed ANOVA, with anthropomorphism as a between-subjects factor and social cognition as a within-subjects factor, revealed a significant main effect of anthropomorphism, $F(1, 276) = 6.45$, $p = 0.012$, $\eta^2 = 0.02$, and a significant main effect of social cognition, $F(1, 276) = 534.31$, $p < 0.001$, $\eta^2 = 0.66$. The interaction between anthropomorphism and social cognition was significant, $F(1, 276) =$

Figure 1

Figure 1: Figure 1

8.00, $p = 0.005$, $\eta^2 = 0.03$. Simple effects analysis showed that the anthropomorphic group ($M = 4.20$, $SD = 0.12$) rated the AI significantly higher on warmth than the non-anthropomorphic group ($M = 3.71$, $SD = 0.12$), $p = 0.003$, $\eta^2 = 0.03$. However, competence ratings did not differ significantly between the anthropomorphic ($M = 5.78$, $SD = 0.06$) and non-anthropomorphic ($M = 5.73$, $SD = 0.06$) groups, $p = 0.58$.

Effect of Anthropomorphism on Cooperation Intention: An independent samples *t*-test with anthropomorphism as the independent variable and cooperation intention as the dependent variable showed no significant difference between the anthropomorphic ($M = 5.12$, $SD = 1.18$) and non-anthropomorphic ($M = 5.00$, $SD = 1.28$) groups, $t(276) = 0.85$, $p = 0.40$, Cohen' s $d = 0.10$.

These results indicate that anthropomorphism significantly affected perceived warmth (partially supporting Hypothesis 2a) but did not directly affect cooperation intention (failing to support Hypothesis 1). We hypothesized that this occurred because perceived warmth-competence is more sensitive to anthropomorphic manipulation, while the manipulation itself, though effective, had a small effect size. Thus, anthropomorphism influences cooperation intention indirectly through perceived warmth. To test this, we conducted a path analysis of anthropomorphism → perceived warmth → cooperation intention.

Path Analysis of Anthropomorphism → Perceived Warmth → Cooperation Intention: Following Hayes (2013), we used Process Macro Model 4 with 5,000 bootstrap samples and a 95% confidence interval. Anthropomorphism group (non-anthropomorphic = 1, anthropomorphic = 2) served as the independent variable, cooperation intention as the dependent variable, and perceived warmth as the mediator. Results showed a significant indirect effect of perceived warmth, with an effect size of 0.22 and a 95% CI excluding zero [0.08, 0.36]. After controlling for perceived warmth, the direct effect of anthropomorphism was non-significant, with a 95% CI of [-0.38, 0.10] containing zero and a direct effect of -0.11.

These results (see Figure 1

) indicate that anthropomorphism affects cooperation intention through the full mediation of perceived warmth, supporting our hypothesis. Another possible reason for the non-significant direct effect is that the performance evaluation scenario might have induced threat perceptions. To rule out this potential confound, Study 1b used a more general anthropomorphic scenario.

2.2 Study 1b

Study 1b employed a more general anthropomorphic scenario to examine the effects of anthropomorphism on AI' s perceived warmth-competence and human-AI cooperation intention.

2.2.1 Participants

Using G*Power 3.1 (Faul et al., 2009) with a medium effect size of* d* = 0.5 (Zhao et al., 2024), $\alpha = 0.05$, and 90% power, at least 172 participants were required. To account for exclusions, we recruited 180 participants (110 female) through credamo, with real-time exclusion of those failing attention checks or completing the survey in under 120 seconds. The mean age was 30.82 years (*SD* = 9.16). All participants provided informed consent and received compensation.

2.2.2 Design and Procedure

Study 1b used the same design and procedure as Study 1a. Participants were randomly assigned to the anthropomorphic or non-anthropomorphic group. The anthropomorphic manipulation materials were as follows:

Anthropomorphic group AI material: "Hi there! My name is Archie, and I' m your intelligent collaboration partner. I' m here to work with you to tackle various tasks and challenges. I will analyze information, generate solutions, predict potential problems, and communicate and coordinate closely with you throughout the process. I will flexibly adjust my strategies based on your feedback and needs to ensure our collaboration runs smoothly and efficiently. Whatever the goal, I am committed to working alongside you to find optimal solutions and accomplish objectives efficiently together. Let' s join hands and leverage our respective strengths!"

Non-anthropomorphic group AI material: "Collaborative Intelligent Tool HAP is a new system designed to assist users in handling various tasks and challenges. This tool analyzes information, generates solutions, predicts potential problems, and maintains communication and coordination with users throughout the process. Collaborative Intelligent Tool HAP adjusts its strategies based on user feedback and needs to ensure smooth and efficient support. Whatever the goal, this tool aims to provide support to users, assist in finding optimal solutions, and help users accomplish objectives efficiently. Collaborative Intelligent Tool HAP is designed to collaborate with users, combining both parties' capabilities."

This manipulation paradigm was adapted from previous anthropomorphism studies (Zhao et al., 2024). To verify effectiveness, participants rated the material' s anthropomorphism level ("Please rate the degree to which it has human characteristics, motivations, intentions, or mental states," Phillips et al., 2018) on a continuous slider scale from 1 (not at all human) to 100 (just like a human).

Cooperation intention was measured as in Study 1a (Cronbach' s $\alpha = 0.81$). Perceived warmth and competence were also measured as before (warmth $\alpha = 0.92$; competence $\alpha = 0.67$).

2.2.3 Results and Discussion

Manipulation Check and Control Variables: The anthropomorphism check showed that the anthropomorphic group ($M = 70.02$, $SD = 18.64$) scored significantly higher than the non-anthropomorphic group ($M = 58.31$, $SD = 22.16$), $t(178) = 3.84$, $p < 0.001$, Cohen' s $d = 0.75$, confirming effective manipulation. Age differences were non-significant, $t(178) = 1.41$, $p = 0.16$. Gender correlated marginally with perceived warmth (r = -0.14, $p = 0.062$), significantly with perceived competence (r = -0.15, $p = 0.043$), and significantly with cooperation intention (r = -0.17, $p = 0.025$), so gender was included as a covariate.

Effect of Anthropomorphism on Perceived Warmth-Competence: A 2 (anthropomorphic/non-anthropomorphic) × 2 (warmth/competence) mixed ANOVA revealed a significant main effect of anthropomorphism, $F(1, 178) = 11.41$, $p = 0.001$, $\eta^2 = 0.06$, and a significant main effect of social cognition, $F(1, 178) = 205.28$, $p < 0.001$, $\eta^2 = 0.54$. The interaction was significant, $F(1, 178) = 23.10$, $p < 0.001$, $\eta^2 = 0.12$. Simple effects analysis showed that the anthropomorphic group ($M = 5.30$, $SD = 0.11$) rated the AI significantly higher on warmth than the non-anthropomorphic group ($M = 4.60$, $SD = 0.11$), $p < 0.001$, $\eta^2 = 0.10$. Competence ratings did not differ significantly between groups ($M = 6.00$ vs. 6.02, $p = 0.75$). These results remained unchanged when controlling for gender.

The non-significant effect of anthropomorphism on perceived competence may be due to ceiling effects caused by the scenario' s emphasis on AI' s high capabilities (mean competence ratings > 6.00).

Effect of Anthropomorphism on Cooperation Intention: Consistent with Study 1a, an independent samples t-test revealed no significant difference in cooperation intention between the anthropomorphic ($M = 6.00$, $SD = 0.67$) and non-anthropomorphic ($M = 5.92$, $SD = 0.72$) groups, $t(178) = 0.78$, $p = 0.44$. Controlling for gender did not change this result, $F(1, 177) = 0.40$, $p = 0.526$.

Path Analysis of Anthropomorphism → Perceived Warmth → Cooperation Intention: Using Process Macro Model 4 with 5,000 bootstrap samples and a 95% CI, we tested perceived warmth as a mediator between anthropomorphism group (non-anthropomorphic = 1, anthropomorphic = 2) and cooperation intention. Results showed a significant indirect effect of perceived warmth, with an effect size of 0.28 and a 95% CI of [0.12, 0.47]. After controlling for perceived warmth, the direct effect of anthropomorphism was non-significant, with a 95% CI of [-0.31, 0.08] containing zero and a direct effect of -0.16. As in Study 1a, these results (see Figure 2 [FIGURE:2]) indicate that anthropo-

morphism affects cooperation intention through the full mediation of perceived warmth. Including gender as a covariate did not alter this conclusion.

Although Study 1' s passive textual manipulation did not reveal a direct effect of anthropomorphism on cooperation intention, it demonstrated that anthropomorphism enhances perceived warmth, which in turn promotes cooperation intention. This supports the core hypothesis that anthropomorphism influences social cognition to facilitate human-AI collaboration. The absence of a direct effect may be attributed to perceived warmth being more sensitive to anthropomorphic manipulation and the relatively small effect size of the passive textual manipulation. Additionally, the scenario description may have influenced cooperation intention. To address these limitations, Study 2 employed an active imagination manipulation paradigm (without scenario descriptions) to further examine anthropomorphism' s effects on cooperation intention and perceived competence, and to systematically test whether both perceived warmth and competence mediate the relationship.

3 Study 2: Active Imagination Manipulation of Anthropomorphism

Study 2 aimed to use active imagination manipulation to investigate the effects of AI anthropomorphism on perceived warmth-competence and cooperation intention, particularly its impact on perceived competence and cooperation intention, and to test whether perceived warmth-competence mediates this relationship.

3.1.1 Participants

G*Power 3.1 (Faul et al., 2009) indicated that for an independent* t*-test with medium effect size* d* = 0.5 (Roesler et al., 2021), $\alpha = 0.05$, and 80% power, at least 128 participants were needed. To account for exclusions, we recruited 136 participants (91 female) through credamo, with real-time exclusion of those failing attention checks or completing the survey in under 120 seconds. The mean age was 30.76 years ($SD = 8.03$). All participants provided informed consent and received compensation.

3.1.2 Design and Procedure

Study 2 used a single-factor between-subjects design with two levels: active imagination high-anthropomorphism group versus low-anthropomorphism group (hereafter high-anthropomorphism and low-anthropomorphism). Participants were randomly assigned and first read the same collaboration scenario as in Study 1.

After answering questions about their reference team (as in Study 1), participants were asked to imagine either a low-anthropomorphism robot (anthropomorphism level 1-33) or a high-anthropomorphism robot (level 67-100). To

control for individual differences in anthropomorphism perception, participants first rated the widely used robotic AI NAO on anthropomorphism (scored A on a 1-100 scale). The NAO image came from the ABOT (Anthropomorphic robot) database (Phillips et al., 2018), with a standard rating of 50. Participants then rated their imagined AI' s anthropomorphism (scored B). This active imagination manipulation, which guides participants to describe or evaluate experimental materials from an anthropomorphic perspective without biasing materials, effectively engages participants' agency and achieves robust manipulation effects (Xu et al., 2017). It also reduces constraints on competence judgments from specific information. Finally, participants reported cooperation intention, perceived warmth, and perceived competence.

Manipulation effectiveness was assessed through two indicators: participants' self-reported anthropomorphism ratings of their imagined robot (using credamo' s slider scale from 1 = "not at all human" to 100 = "just like a human," Phillips et al., 2018) and standardized anthropomorphism scores (X1) to prevent interference from the instructional ranges.

Cooperation intention was measured as in Study 1 (Cronbach' s $\alpha = 0.92$). Perceived warmth and competence were measured as before (warmth $\alpha = 0.92$; competence $\alpha = 0.93$).

3.2 Results

Manipulation Check and Control Variables: Independent samples t-tests confirmed that high-anthropomorphism participants rated their imagined robot as significantly more anthropomorphic ($M = 72.99$, $SD = 18.20$) than low-anthropomorphism participants ($M = 27.64$, $SD = 17.07$), $t(134) = 14.98$, $p < 0.001$, Cohen' s $d = 2.57$. Standardized anthropomorphism scores also differed significantly ($M = 336.21$ vs. 49.16), $t(134) = 3.01$, $p = 0.003$, Cohen' s $d = 0.52$, confirming effective manipulation.

Ratings of the NAO robot did not differ between high- ($M = 43.28$, $SD = 28.42$) and low-anthropomorphism ($M = 44.40$, $SD = 23.92$) groups, $t(134) = 0.25$, $p = 0.80$, ruling out individual differences in anthropomorphism perception as a confound. Age differences were non-significant, $t(134) = 1.78$, $p = 0.077$. Gender was not significantly correlated with perceived warmth ($r = -0.003$, $p = 0.975$), competence ($r = 0.02$, $p = 0.795$), or cooperation intention ($r = -0.10$, $p = 0.259$), so it was excluded as a covariate.

Effect of Anthropomorphism on Cooperation Intention: An independent samples t-test revealed that high-anthropomorphism participants reported significantly higher cooperation intention ($M = 5.91$, $SD = 0.82$) than low-anthropomorphism participants ($M = 5.06$, $SD = 1.43$), $t(134) = 4.31$, $p < 0.001$, Cohen' s $d = 0.74$ (see Figure 3 [FIGURE:3]).

Effect of Anthropomorphism on Perceived Warmth-Competence: A 2 (high/low anthropomorphism) × 2 (warmth/competence) mixed

ANOVA showed a significant main effect of anthropomorphism, with high-anthropomorphism (M = 5.60, SD = 0.13) scoring higher than low-anthropomorphism (M = 4.31, SD = 0.13), $F(1, 134) = 48.86$, $p < 0.001$, $\eta^2 = 0.27$. The main effect of social cognition was also significant, with perceived competence (M = 5.35, SD = 0.10) higher than perceived warmth (M = 4.56, SD = 0.11), $F(1, 134) = 70.26$, $p < 0.001$, $\eta^2 = 0.34$. The interaction between anthropomorphism and social cognition was non-significant (see Figure 4 [FIGURE:4]), $F(1, 134) = 0.87$, $p = 0.35$.

Mediating Role of Perceived Warmth-Competence: To test whether perceived warmth and competence mediate the effect of anthropomorphism on cooperation intention, we used Process Macro Model 4 with 5,000 bootstrap samples and a 95% CI. Anthropomorphism group (low = 1, high = 2) was the independent variable, cooperation intention the dependent variable, and perceived competence and warmth parallel mediators. Results showed significant indirect effects for both competence (effect = 0.59, 95% CI [0.32, 0.87]) and warmth (effect = 0.23, 95% CI [0.03, 0.45]). After controlling for both mediators, the direct effect of anthropomorphism was non-significant (95% CI [-0.46, 0.16], direct effect = -0.12). These results (see Figure 5 [FIGURE:5]) indicate that perceived competence and warmth fully mediate the effect of anthropomorphism on cooperation intention.

3.3 Discussion

Study 2 found that anthropomorphism enhances perceived warmth and competence, increases cooperation intention, and that both dimensions fully mediate this relationship, fully supporting our core hypotheses. The active imagination manipulation addressed limitations of Study 1' s passive textual approach, demonstrating that anthropomorphism improves not only perceived warmth but also competence and cooperation intention.

However, Study 2' s active imagination manipulation had limitations. While we conducted manipulation checks, the lack of control and verification of participants' imagined content could render the manipulation ineffective. Therefore, Study 3 builds on Study 2 by adding active textual descriptions to supplement imagination, addressing the lack of content control while enhancing manipulation strength—multiple anthropomorphic methods have been shown to produce stronger effects (e.g., Puzakova et al., 2013; Wan et al., 2017). Additionally, Study 3 examines the moderating role of realistic threat on anthropomorphism' s effects on perceived warmth-competence and cooperation intention.

4 Study 3: Moderating Role of Realistic Threat

Study 3 had two objectives: (1) to replicate the effects of active imagination anthropomorphism (supplemented with textual descriptions) on perceived warmth-competence and cooperation intention, and test the mediating role of these dimensions; and (2) to examine the moderating effect of realistic threat.

4.1.1 Participants

G*Power 3.1 (Faul et al., 2009) indicated that for an independent* t*-test with medium effect size* d* = 0.5 (Roesler et al., 2021), $\alpha = 0.05$, and 90% power, at least 172 participants were needed. To account for exclusions, we recruited 178 participants (114 female) through credamo, with real-time exclusion of those failing attention checks or taking over 2000 seconds, using rolling recruitment. The mean age was 30.97 years ($SD = 9.07$). All participants provided informed consent and received compensation.

4.1.2 Design and Procedure

Study 3a was a preregistered study (https://aspredicted.org/d3ng-h7qz.pdf) using a single-factor between-subjects design with two levels: low-anthropomorphism versus high-anthropomorphism groups (hereafter low- and high-anthropomorphism). Participants were randomly assigned. To control for individual differences in anthropomorphic tendencies, all participants completed the Anthropomorphism Questionnaire (AQ; Neave et al., 2015) before the experiment, using five items about electronic products (e.g., "Sometimes I think my computer runs slower on purpose after I yell at it") on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Participants then read the collaboration scenario and imagined/described either a low- or high-anthropomorphism AI.

High-anthropomorphism AI material: "Please imagine a highly anthropomorphic robotic AI that will serve as your collaboration partner. This robot possesses near-human characteristics, motivations, intentions, or mental states." Participants were asked to give it a human name (e.g., Archie) and describe its appearance/movements (e.g., subtle facial expressions, body language, eye contact), personality/emotions (e.g., understanding your emotions, showing humor, empathy, or personality), and behavioral patterns (e.g., proactive communication, flexible adaptation to unexpected situations) in 2-3 sentences each.

Low-anthropomorphism AI material: "Please imagine a low-anthropomorphism robotic AI that will serve as your collaborative tool. This robot has basic functions." Participants gave it a tool-like name (e.g., XR-7, Collaborator-9) and described its appearance/movements (e.g., mechanical appearance, programmed movements, no human-like expressions), personality/emotions (e.g., operates strictly by command, no personality, emotions, or autonomous intentions), and behavioral patterns (e.g., passively responds to needs, communicates through data/indicator lights) in 2-3 sentences each.

This paradigm was adapted from previous research (e.g., Aggarwal & McGill, 2012; Kim & Kramer, 2015). A pilot test ($N = 60$) with four psychology students rating the three anthropomorphic dimensions (Kendall' s $W = 0.91$, $p < 0.001$) confirmed the manipulation' s effectiveness: high-anthropomorphism ($M = 4.76$, $SD = 1.03$) scored significantly higher than low-anthropomorphism ($M = 1.73$, $SD = 0.99$), $t(58) = 11.69$, $p < 0.001$, Cohen' s $d = 3.02$.

After this task, participants rated their described robot' s anthropomorphism ("Please rate the degree to which it has human characteristics, motivations, intentions, or mental states," Phillips et al., 2018). They then read a news article about low realistic threat from robots (see Appendix; Jackson et al., 2020; Xu et al., 2024) and completed a realistic threat perception check (Yogeeswaran et al., 2016): (1) "Increasing robot use in daily life is causing human unemployment" ; (2) "Robots cannot replace people' s jobs" (reverse-scored); (3) "In the long run, robots pose a direct threat to human safety and well-being" ; (4) "Robot technology development threatens human employment and opportunities" ; (5) "Increasing prevalence of robots in daily life threatens human safety." Cronbach' s α was 0.83.

We also measured perceived in-group status using a self-developed item ("Do you view the robot 'XXX' you just imagined as an in-group team member?") and an adapted Inclusion of Other in the Self scale (Aron et al., 1992), both on 7-point scales. Before measurement, participants read an explanation of in-group/out-group concepts in social psychology. Cronbach' s α was 0.81.

Cooperation intention was measured as in Study 1 (Cronbach' s $\alpha = 0.88$). Perceived warmth and competence were measured as before (warmth $\alpha = 0.96$; competence $\alpha = 0.89$).

4.1.3 Results and Analysis

Manipulation Check and Control Variables: The high-anthropomorphism group ($M = 78.96$, $SD = 11.51$) scored significantly higher on anthropomorphism than the low-anthropomorphism group ($M = 29.03$, $SD = 26.66$), $t(176) = 16.22$, $p < 0.001$, Cohen' s $d = 2.43$, confirming effective manipulation. Age differences were non-significant, $t(176) = 1.53$, $p = 0.13$, as were individual differences in anthropomorphic tendencies, $t(176) = 0.05$, $p = 0.96$. Gender was not significantly correlated with perceived warmth ($r = 0.03$, $p = 0.737$), competence ($r = 0.05$, $p = 0.476$), or cooperation intention ($r = 0.02$, $p = 0.808$), so it was excluded as a covariate.

Effect of Anthropomorphism on Perceived Warmth-Competence: A 2 (high/low anthropomorphism) × 2 (warmth/competence) mixed ANOVA, with anthropomorphism as a between-subjects factor, social cognition as a within-subjects factor, and AQ scores as a covariate, revealed a significant main effect of anthropomorphism, $F(1, 175) = 81.92$, $p < 0.001$, $\eta^2 = 0.32$, and a significant main effect of social cognition, $F(1, 175) = 21.61$, $p < 0.001$, $\eta^2 = 0.11$. The interaction was significant (see Figure 6 [FIGURE:6]), $F(1, 175) = 64.04$, $p < 0.001$, $\eta^2 = 0.27$.

Simple effects analysis showed that the high-anthropomorphism group ($M = 5.41$, $SD = 0.14$) rated the AI significantly higher on warmth than the low-anthropomorphism group ($M = 3.36$, $SD = 0.14$), $p < 0.001$, $\eta^2 = 0.38$. The high-anthropomorphism group ($M = 5.80$, $SD = 0.10$) also rated competence significantly higher than the low-anthropomorphism group ($M = 5.19$, $SD =$

0.10), $p < 0.001$, $\eta^2 = 0.10$. These results indicate that anthropomorphism enhances both perceived warmth and competence, with a larger effect on warmth.

Effect of Anthropomorphism on Cooperation Intention: An independent samples t-test showed that the high-anthropomorphism group ($M = 5.70$, $SD = 0.85$) reported significantly higher cooperation intention than the low-anthropomorphism group ($M = 5.01$, $SD = 1.14$), $t(176) = 4.64$, $p < 0.001$, Cohen' s $d = 0.70$ (see Figure 8 [FIGURE:8]). This difference remained significant after controlling for AQ scores, $F(1, 175) = 21.47$, $p < 0.001$, $\eta^2 = 0.11$.

Mediating Role of Perceived Warmth-Competence: Using Process Macro Model 4 with 5,000 bootstrap samples and a 95% CI, we tested parallel mediation with anthropomorphism group (low = 1, high = 2) as the independent variable, cooperation intention as the dependent variable, and perceived competence and warmth as mediators, controlling for AQ scores. Results showed significant indirect effects for both competence (effect = 0.28, 95% CI [0.12, 0.48]) and warmth (effect = 0.38, 95% CI [0.14, 0.61]). After controlling for both mediators, the direct effect of anthropomorphism was non-significant (95% CI [-0.29, 0.31], direct effect = 0.007). These results (see Figure 9 [FIGURE:9]) indicate that perceived warmth and competence fully mediate the effect of anthropomorphism on cooperation intention.

Alternative Mediation via Perceived Threat: To rule out perceived threat as an alternative mediator, we conducted a simple mediation analysis with anthropomorphism group as the independent variable, perceived realistic threat as the mediator, and cooperation intention as the dependent variable, controlling for AQ scores. The indirect effect of perceived realistic threat was non-significant (effect = -0.03, 95% CI [-0.11, 0.04]).

Study 3a demonstrated that under low realistic threat, active imagination manipulation of anthropomorphism enhances perceived warmth and competence and increases cooperation intention, with both dimensions fully mediating this effect—consistent with Study 2 and fully supporting our core hypotheses. The study also ruled out alternative mediation through perceived realistic threat.

4.2.1 Participants

For Study 3b, G*Power 3.1 indicated that 172 participants were needed for 90% power. We recruited 178 participants (109 female) through credamo, excluding those failing attention checks or taking over 2000 seconds. Mean age was 31.25 years (*SD* = 9.22). All provided informed consent and received compensation.

4.2.2 Design and Procedure

Study 3b was preregistered (https://aspredicted.org/d3ng-h7qz.pdf) with the same design as Study 3a. Participants completed the AQ (Neave et al., 2015)

and imagined/described low- or high-anthropomorphism AI using the same materials as Study 3a.

Afterward, participants rated their robot' s anthropomorphism as before. They then read a news article about high realistic threat from robots (see Appendix; Jackson et al., 2020; Xu et al., 2024) and completed the realistic threat perception check (Yogeeswaran et al., 2016; Cronbach' s $\alpha = 0.80$). Perceived in-group status was measured as in Study 3a (Cronbach' s $\alpha = 0.87$). Cooperation intention (Cronbach' s $\alpha = 0.91$), perceived warmth ($\alpha = 0.96$), and competence ($\alpha = 0.86$) were measured as before.

4.2.3 Results and Analysis

Manipulation Check and Control Variables: The high-anthropomorphism group ($M = 81.48$, $SD = 10.11$) scored significantly higher on anthropomorphism than the low-anthropomorphism group ($M = 30.35$, $SD = 26.37$), $t(176) = 17.08$, $p < 0.001$, Cohen' s $d = 2.56$, confirming effective manipulation. Age differences were non-significant, $t(176) = 1.00$, $p = 0.32$, as were AQ differences, $t(176) = 0.64$, $p = 0.39$. Gender was not significantly correlated with perceived warmth ($r = 0.03$, $p = 0.714$), competence ($r = 0.08$, $p = 0.314$), or cooperation intention ($r = 0.03$, $p = 0.745$), so it was excluded as a covariate.

Effect of Anthropomorphism on Perceived Warmth-Competence: A 2 (high/low anthropomorphism) × 2 (warmth/competence) mixed ANOVA, controlling for AQ scores, revealed a significant main effect of anthropomorphism, $F(1, 175) = 68.95$, $p < 0.001$, $\eta^2 = 0.28$, and a significant main effect of social cognition, $F(1, 175) = 39.90$, $p < 0.001$, $\eta^2 = 0.19$. The interaction was significant (see Figure 7 [FIGURE:7]), $F(1, 175) = 57.47$, $p < 0.001$, $\eta^2 = 0.25$.

Simple effects analysis showed that the high-anthropomorphism group ($M = 5.44$, $SD = 0.14$) rated the AI significantly higher on warmth than the low-anthropomorphism group ($M = 3.58$, $SD = 0.14$), $p < 0.001$, $\eta^2 = 0.33$. The high-anthropomorphism group ($M = 5.91$, $SD = 0.09$) also rated competence significantly higher than the low-anthropomorphism group ($M = 5.32$, $SD = 0.09$), $p < 0.001$, $\eta^2 = 0.11$. These results indicate that anthropomorphism enhances both perceived warmth and competence, with a larger effect on warmth.

Effect of Anthropomorphism on Cooperation Intention: An independent samples t-test showed that the high-anthropomorphism group ($M = 5.69$, $SD = 0.88$) reported significantly higher cooperation intention than the low-anthropomorphism group ($M = 5.10$, $SD = 1.10$), $t(176) = 3.97$, $p < 0.001$, Cohen' s $d = 0.60$ (see Figure 8). This difference remained significant after controlling for AQ scores, $F(1, 175) = 15.64$, $p < 0.001$, $\eta^2 = 0.08$.

Mediating Role of Perceived Warmth-Competence: Parallel mediation analysis using Process Macro Model 4 (5,000 bootstrap samples, 95% CI), controlling for AQ scores, showed significant indirect effects for both competence (effect = 0.35, 95% CI [0.19, 0.54]) and warmth (effect = 0.33, 95% CI [0.12,

0.55]). After controlling for both mediators, the direct effect of anthropomorphism was non-significant (95% CI [-0.38, 0.14], direct effect = -0.12). These results (see Figure 10 [FIGURE:10]) indicate that perceived warmth and competence fully mediate the effect of anthropomorphism on cooperation intention.

Study 3b demonstrated that under high realistic threat, active imagination manipulation of anthropomorphism still enhances perceived warmth and competence and increases cooperation intention, with both dimensions serving as mediators—consistent with Studies 2 and 3a and supporting our core hypotheses.

4.3 Realistic Threat Moderation Analysis

To test whether realistic threat moderates anthropomorphism' s effects, we combined data from Studies 3a and 3b.

Effects of Anthropomorphism and Realistic Threat on Perceived Warmth-Competence: An independent samples t-test confirmed that the high realistic threat group ($M = 4.24$, $SD = 1.10$) perceived significantly higher threat than the low threat group ($M = 3.15$, $SD = 1.74$), $t(354) = 9.09$, $p < 0.001$, Cohen' s $d = 0.75$, confirming effective threat manipulation.

A 2 (high/low anthropomorphism) $\times$ 2 (high/low realistic threat) ANOVA on perceived warmth, controlling for AQ scores, showed a significant main effect of anthropomorphism, $F(1, 355) = 192.66$, $p < 0.001$, $\eta^2 = 0.35$, but no main effect of realistic threat, $F(1, 355) = 0.71$, $p = 0.40$, and no significant interaction, $F(1, 355) = 0.78$, $p = 0.51$.

A similar ANOVA on perceived competence revealed a significant main effect of anthropomorphism, $F(1, 355) = 31.91$, $p < 0.001$, $\eta^2 = 0.10$, but no main effect of realistic threat, $F(1, 355) = 1.49$, $p = 0.22$, and no significant interaction, $F(1, 355) = 0.01$, $p = 0.90$.

These results indicate that anthropomorphism enhances perceived warmth and competence, while realistic threat has no significant effect on these perceptions.

Effects of Anthropomorphism and Realistic Threat on Cooperation Intention: A 2 $\times$ 2 ANOVA on cooperation intention, controlling for AQ scores, showed a significant main effect of anthropomorphism, $F(1, 355) = 37.23$, $p < 0.001$, $\eta^2 = 0.10$, but no main effect of realistic threat, $F(1, 355) = 0.13$, $p = 0.72$, and no significant interaction, $F(1, 355) = 0.23$, $p = 0.63$.

These results indicate that anthropomorphism promotes cooperation intention, while realistic threat has no significant effect.

4.4 Discussion

Combined results from Studies 3a and 3b failed to support Hypothesis 3 regarding realistic threat moderation. We reason that this occurred because realistic threat does not affect perceived warmth-competence, and more fundamentally,

because realistic threat does not influence in-group/out-group perception. Comparing in-group perception scores across high and low realistic threat conditions revealed no significant difference (M = 4.04, SD = 1.48 vs. M = 3.97, SD = 1.43), $t(354)$ = -0.42, $p = 0.68$. Since threat moderation fundamentally operates through out-group perception triggered by threat, and previous research found uniqueness threat more influential than realistic threat in human-robot contexts (Huang et al., 2021; Xu et al., 2024), Study 4 examines the moderating role of uniqueness threat. Additionally, to enhance ecological validity, Study 4 uses more intuitive and visual robotic AI appearance manipulation.

5 Study 4: Appearance Manipulation of Anthropomorphism

Study 4 used robotic AI appearance manipulation to examine anthropomorphism' s effects on perceived warmth-competence and cooperation intention, and further explore the moderating role of uniqueness threat.

5.1.1 Participants

G*Power 3.1 indicated that 172 participants were needed for 90% power. We recruited 178 participants (126 female) through credamo, excluding those failing attention checks or completing the survey in under 120 seconds. Mean age was 30.33 years (*SD* = 7.99). All provided informed consent and received compensation.

5.1.2 Design and Procedure

Study 4 used a single-factor between-subjects design with two levels: high- versus low-anthropomorphism groups. After reading the collaboration scenario and answering reference team questions, participants were randomly presented with robotic AI images and evaluated the AI' s anthropomorphism level, cooperation intention, and perceived warmth-competence. They then completed individual difference measures of uniqueness threat and reported demographics.

We adapted Ferrari et al.' s (2016) paradigm using robotic images from the standardized Anthropomorphic roBOT (ABOT) database (Phillips et al., 2018), widely used in anthropomorphism research (Roesler et al., 2021). We selected 12 robotic images: six low-anthropomorphism (standardized scores 1-33) and six high-anthropomorphism (67-100). Manipulation checks used subjective anthropomorphism ratings (Phillips et al., 2018).

Cooperation intention (Cronbach' s α = 0.91), perceived warmth (α = 0.90), and competence (α = 0.92) were measured as before. Individual differences in uniqueness threat were measured using Yogeeswaran et al.' s (2016) uniqueness threat subscale: (1) "Widespread robot use in daily life bothers me because it blurs the line between humans and machines" ; (2) "Lifelike robots are unsettling because they' re almost indistinguishable from humans" ; (3) "Recent

technological advances challenge human essence" ; (4) "Robotic technological progress threatens human uniqueness" ; (5) "Robots are blurring boundaries between humans and machines." This 7-point Likert scale had Cronbach' s α = 0.87.

5.2 Results and Analysis

Manipulation Check and Control Variables: The high-anthropomorphism group (M = 56.77, SD = 25.52) scored significantly higher on anthropomorphism than the low-anthropomorphism group (M = 23.37, SD = 20.22), $t(176)$ = 9.64, $p < 0.001$, Cohen' s d = 1.45, confirming effective manipulation. Age differences were non-significant, $t(176)$ = 0.45, p = 0.65. Gender was not significantly correlated with competence (r = -0.003, p = 0.973) or cooperation intention (r = -0.07, p = 0.385), but correlated significantly with warmth (r = -0.19, p = 0.013), so gender was included as a covariate.

Effect of Appearance Anthropomorphism on Cooperation Intention: An independent samples t-test showed a marginally significant difference, with the high-anthropomorphism group (M = 4.70, SD = 1.47) reporting lower cooperation intention than the low-anthropomorphism group (M = 5.10, SD = 1.06), $t(176)$ = -1.95, p = 0.052, Cohen' s d = 0.29. After controlling for gender, the difference became significant, $F(1, 175)$ = 4.34, p = 0.039, η^2 = 0.02. Contrary to hypotheses, high anthropomorphism reduced cooperation intention.

Effect of Appearance Anthropomorphism on Perceived Warmth-Competence: A 2 (high/low anthropomorphism) × 2 (warmth/competence) mixed ANOVA showed no significant main effect of anthropomorphism, $F(1, 176)$ = 0.15, p = 0.70, but a significant main effect of social cognition, with competence (M = 5.11, SD = 0.10) rated higher than warmth (M = 4.19, SD = 0.10), $F(1, 176)$ = 105.85, $p < 0.001$, η^2 = 0.38. The interaction was non-significant, $F(1, 176)$ = 0.36, p = 0.56. These results remained unchanged when controlling for gender.

These results indicate that appearance anthropomorphism does not affect perceived warmth or competence.

Mediating Role of Perceived Warmth-Competence in Appearance Anthropomorphism: Parallel mediation analysis using Process Macro Model 4 (5,000 bootstrap samples, 95% CI) with appearance anthropomorphism group (low = 1, high = 2) as the independent variable, cooperation intention as the dependent variable, and perceived competence and warmth as mediators showed non-significant indirect effects for both competence (95% CI [-0.25, 0.12]) and warmth (95% CI [-0.09, 0.07]). After controlling for both mediators, the direct effect of appearance anthropomorphism was significant (95% CI [-0.53, -0.07]), indicating that appearance anthropomorphism directly and negatively affects cooperation intention without mediation through warmth-competence (see Figure 11 [FIGURE:11]).

Moderating Role of Uniqueness Threat on Appearance Anthropomorphism' s Effect on Perceived Warmth-Competence: Using Process Macro Model 1, we tested uniqueness threat as a moderator of the effect of appearance anthropomorphism group on perceived warmth. Results showed non-significant main effects of anthropomorphism (b = -0.01, SE = 0.19, p = 0.06, 95% CI [-0.40, 0.38]) and uniqueness threat (b = 0.32, SE = 0.24, p = 0.18, 95% CI [-0.15, 0.79]), and a non-significant interaction (b = -0.27, SE = 0.15, p = 0.07, 95% CI [-0.56, 0.02]). Results remained unchanged when controlling for gender.

A similar analysis on perceived competence also showed non-significant main and interaction effects, with the uniqueness threat interaction remaining non-significant (b = -0.18, SE = 0.13, p = 0.17, 95% CI [-0.44, 0.08]).

Moderating Role of Uniqueness Threat on Appearance Anthropomorphism' s Effect on Cooperation Intention: Moderation analysis using Process Macro Model 1 revealed a significant main effect of appearance anthropomorphism (b = -0.38, SE = 0.19, p = 0.04, 95% CI [-0.75, -0.01]), a non-significant main effect of uniqueness threat (b = 0.31, SE = 0.22, p = 0.16, 95% CI [-0.13, 0.76]), and a significant interaction (b = -0.33, SE = 0.14, p = 0.02, 95% CI [-0.61, -0.06]) (see Figure 12 [FIGURE:12]). Simple slopes analysis indicated that under low uniqueness threat, anthropomorphism groups did not differ significantly in cooperation intention (b = 0.07, SE = 0.26, p = 0.79, 95% CI [-0.45, 0.59]). However, under high uniqueness threat, the groups differed significantly (b = -0.83, SE = 0.26, p = 0.002, 95% CI [-1.35, -0.31]). Results remained significant when controlling for gender.

These findings indicate that individual differences in uniqueness threat moderate the effect of appearance anthropomorphism on cooperation intention. For individuals low in uniqueness threat, anthropomorphism has no significant effect on cooperation intention. For those high in uniqueness threat, anthropomorphism negatively affects cooperation intention—high anthropomorphism reduces cooperation intention compared to low anthropomorphism.

This seemingly contradictory result may stem from Study 4' s use of ABOT database ratings (based solely on physical appearance) to group participants, whereas our study measured subjective psychological anthropomorphism (degree of human characteristics, motivations, intentions, or mental states). To test this explanation, we reanalyzed the data using subjective anthropomorphism ratings as the independent variable.

5.3 Discussion

Study 4 found that appearance anthropomorphism marginally reduced cooperation intention, with high anthropomorphism groups showing lower intention than low anthropomorphism groups. Uniqueness threat moderated this effect: under low threat, anthropomorphism had no effect; under high threat, it negatively affected cooperation intention. Appearance anthropomorphism

directly influenced cooperation intention without mediating through warmth-competence.

These results diverge from previous studies but align with the uncanny valley effect, where highly human-like but non-human physical appearance can trigger discomfort and threat (Mori, 1970). The moderation analysis suggests that uniqueness threat explains why appearance anthropomorphism reduces cooperation intention: for high-threat individuals, more human-like appearance increases threat perception, reducing cooperation intention. For low-threat individuals, appearance anthropomorphism does not affect cooperation intention.

Notably, supplementary analysis using subjective anthropomorphism ratings yielded results consistent with our core hypothesis: subjective anthropomorphism significantly promoted cooperation intention through enhanced warmth-competence (see Figure 13 [FIGURE:13]). This suggests that the inconsistency arises from different operational definitions: appearance anthropomorphism refers to physical similarity (ABOT' s "how human does this robot look physically?"), while subjective anthropomorphism refers to psychological perception of human characteristics, motivations, intentions, or mental states (Epley et al., 2007). The effective anthropomorphism that promotes cooperation is psychological, not physical. Meta-analytic evidence supports this: appearance anthropomorphism shows small effects on human-robot interaction ($d = 0.5$, Roesler et al., 2021), while subjective psychological anthrop