Human-AI Security Trust in Smart Home Robots and Its Influencing Factors

Authors: Shanshan You, Qi Yue, Chen Junting, Luo Lei, Zhang Kan, Qi Yue

Date: 2025-07-07T12:49:26+00:00

Abstract

With the development of smart home robot technology, security risks have emerged as a new challenge for human-robot trust. This study proposes and validates a new dimension of trust in smart home robots—security trust. To this end, Study 1 developed a security trust scale for smart home robots and verified the stability, reliability, and validity of the three-factor structure of human-robot trust. Studies 2 and 3 conducted in-depth analyses of the impact of robots' static and dynamic features on security trust among human and artificial intelligence (AI) users. The results revealed that, regarding static features, people exhibited higher levels of security trust in robots with shorter heights and less conspicuous cameras; furthermore, the degree of robot anthropomorphism influenced human sensitivity to these static features. Regarding dynamic features, slower robot movement speed and camera shutdown actions enhanced human security trust, while the influence of these dynamic features varied across different scenarios. Additionally, AI and humans demonstrated a certain degree of consistency in security trust, but overall, AI exhibited lower sensitivity to robot cameras than humans. The findings of this study provide important theoretical support and practical guidance for the design and manufacturing of home robots.

Full Text

Safety Trust in Intelligent Domestic Robots: Human and AI Perspectives on Trust and Relevant Influencing Factors

YOU Shanshan[1,2], **QI Yue**[1,2], **CHEN Junting**[1,2], **LUO Lei**[1,2], **ZHANG Kan**[3],

[1] Department of Psychology, Renmin University of China, Beijing 100872, China

[2] Laboratory of the Department of Psychology, Renmin University of China, Beijing 100872, China

[3] State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract

As smart home robot technology advances, safety concerns have emerged as a new challenge in human-robot trust. This study proposes and validates a novel dimension of trust for smart home robots—safety trust. Study 1 developed a safety trust scale for smart home robots and confirmed the stability and validity of a three-factor structure of human-robot trust. Studies 2 and 3 conducted in-depth analyses of how static and dynamic robot features affect safety trust among human and AI users. The findings reveal that, regarding static features, people exhibit higher safety trust toward shorter robots with less conspicuous cameras; furthermore, the degree of robot anthropomorphism influences human sensitivity to these static features. Regarding dynamic features, slower robot movement speeds and camera-off actions enhance human safety trust, with these effects varying across different scenarios. Additionally, AI demonstrates some consistency with humans in safety trust, though overall AI shows lower sensitivity to robot cameras than humans do. These results provide important theoretical support and practical guidance for the design and manufacturing of home robots.

Keywords: human-robot trust, safety trust, smart home robots, usage intention, large language models

In the era of Artificial Intelligence (AI), robots have become important participants in human social activities, playing significant roles in education [?, ?], healthcare [?, ?], commerce [?, ?], and other domains. In household contexts, the application of smart home robots is gradually becoming widespread [?, ?]. For instance, the robot Kuavo demonstrated laundry and gardening functions in home environments at the 2024 China Appliance and Electronics World Expo (AWE) [?, ?]. Beyond completing chores like room cleaning, robots such as Jibo can also provide emotional companionship and support to family members through sensing and interaction technologies [?, ?]. As technology progresses, robots' intelligence levels continue to improve, and the relationship between humans and robots has shifted from a tool-based human-machine relationship to a social relationship between humans and intelligent agents [?, ?, ?]. Human trust in robots has also evolved accordingly, gradually approaching interpersonal trust [?, ?, ?]. In human-robot interaction, user trust is a critical determinant of whether people will use robots [?, ?].

Research on trust in smart home robots not only facilitates their application and promotion but also represents an important topic in studying human-machine relationships in the intelligent era. The significant differences between AI and traditional technologies have generated many new concerns, deepening the complexity of human-AI interaction. For example, issues such as personality perception of intelligent systems (e.g., personalized characteristics of AI), relational connection and teamwork between humans and intelligent systems have triggered new explorations in human-AI interaction [?, ?, ?, ?, ?]. To adapt to increasingly complex human-AI interaction environments, trust research needs further deepening. In the specific application scenario of smart homes, intelligent robots' autonomous activities are closely related to family members' daily lives and privacy protection, making high robot intelligence potentially trigger new user concerns. Although previous research has confirmed the importance of safety perception in human-robot interaction, it has not deeply explored the potential crisis that safety perception may bring to human-robot trust [?, ?, ?]. Therefore, this study proposes and validates a new dimension of trust for smart home robots—safety trust. From a theoretical perspective, the proposal of safety trust expands and improves existing trust models, providing a new perspective for trust research in the new era of human-AI interaction. From a practical perspective, in-depth research on safety trust can provide valuable guidance for the design and optimization of smart home robots, thereby promoting their widespread application in family life.

Meanwhile, AI' s role has gradually expanded from traditional task executors to decision-making agents. Particularly in multi-agent collaborative systems, AI not only needs to take responsibility as a task executor but also needs to participate in decision-making processes as a trustor [?, ?, ?]. This change requires AI not only to understand and respond to human trust expectations but also to demonstrate consistency and transparency in decision-making processes. Therefore, exploring AI's trust patterns, especially trust mechanisms in complex home environments, helps us better understand and design deeper interaction methods for human-AI cooperation, and provides theoretical support for future applications of smart home robots in more scenarios.

1.1 Smart Home Robots

Smart home robots belong to a category of service robots that provide services to users in home environments, possessing certain perception, interaction, and learning capabilities [?, ?]. Different from early cleaning robots, today' s home robots have gradually achieved diversified functions [?, ?]. For example, the home robot designed by Kao and Wang (2015) has functions including conversational interaction, photography, remote monitoring, and timed reminders. The companion robot Buddy can patrol the house, alert when discovering hazards, and integrate with other smart home products to gain remote control of appliances [?, ?]. Currently, smart home robots have diverse functions that continue to evolve with technological progress, making it necessary to define the scope of this study. Lee (2021) conducted a systematic literature review of the service robot field, classifying service robots into four types (professional non-social, professional social, home/personal non-social, and home/personal social) and summarizing key technical areas (grasping, detection, navigation, human-robot interaction, and architecture/platform, etc.). Based on this, the smart home robots discussed in this paper belong to the home/personal social service robot category, used in home scenarios, possessing grasping, detection, navigation, and information storage functions, capable of interacting with users, and presented as independent architectures rather than platforms [?, ?]. As smart home robots increasingly enter ordinary families, studying user trust in them will lay a theoretical foundation for their promotion and application.

1.2 Human Safety Trust in Smart Home Robots

Human-robot trust is an attitude that an agent can help individuals achieve goals under conditions of known uncertainty and vulnerability [?, ?]. Billings et al. reviewed 302 definitions of trust, including 220 interpersonal trust definitions and 82 automation trust definitions, finding that numerous automation trust definitions involved features such as user expectations of automated systems, confidence, risk, vulnerability, dependence, attitude, and cooperation [?, ?]. These definitions reveal the core characteristics of human-robot trust. Although people rarely report in qualitative studies that they choose to use robots because of trust factors, human-robot trust has been confirmed in multiple studies to affect people' s choice to use robots [?, ?, ?]. As technology continues to advance, the connotation of human-robot trust has gradually enriched from trust in basic tool use (e.g., mechanical equipment) to trust in systems with certain autonomy (e.g., navigation systems), and now to trust in AI systems with complex interaction capabilities, with dimensions progressively enriched and deepened [?, ?].

Human-robot trust has some similarities with interpersonal trust, such as both involving the trustor' s evaluation of the trustee' s reliability, motivation, and capability [?, ?, ?], but they still show significant differences in multiple aspects. Interpersonal trust is usually based on emotional connection and social interaction experience, while human-robot trust relies more on rational evaluation of technical functions. Additionally, interpersonal trust is bidirectional, while human-robot trust is more unidirectional, representing human trust in machines. However, as machine intelligence levels develop, trust between humans and intelligent machines gradually changes, with human-robot trust increasingly approaching interpersonal trust, and many studies have begun exploring human trust patterns in AI from an interpersonal trust perspective [?, ?, ?, ?, ?].

The connotation of human-robot trust has gradually enriched with the development of human-robot interaction technology [?, ?]. In the initial stage, human trust in robots was mainly based on performance levels, meaning people developed trust in robots based on their capability and reliability in completing tasks [?, ?, ?, ?]. As robot interactivity improved, people gradually recognized that beyond performance, the traits and states robots exhibited during interaction also played important roles in trust building [?, ?]. For example, appearance cues such as robots' facial width-to-height ratio, gender characteristics, and anthropomorphism level can activate people' s emotional responses, thereby affecting trust [?, ?, ?, ?]. This trust dimension based on emotional relationships stems from users viewing robots as objects of social interaction, developing emotional trust based on the care and concern robots demonstrate [?, ?, ?]. In scale development, although existing human-robot trust scales [?, ?, ?] contain items reflecting performance trust and relational trust, few scales explicitly distinguish between these two dimensions.

With the rise of cloud computing technology and improvements in robot intelligence levels, human-robot interaction technology will inevitably further expand the evaluation of human-robot trust. Recent research on human interaction with intelligent robots has found that intelligent robots pose numerous safety risks, such as privacy leakage and personal safety risks [?, ?]. On December 26, 2023, the UK' s *Daily Mail* reported an incident two years prior where a robot "attacked" an engineer at a Tesla factory. Söderlund (2023) also emphasized that service robots can easily generate privacy leakage issues, which reduces people' s overall evaluation of robots. Such intentional or unintentional personal or privacy safety risks may not cause individuals to worry about robot performance, nor affect whether individuals view robots as social interaction objects, but they are likely to reduce individuals' trust in robots due to safety risk concerns. Especially in home environments, people' s concerns about privacy and security are more prominent [?, ?]. Therefore, this study proposes that for smart home robots, human-robot trust has generated a new dimension: a safety-based dimension, meaning people develop trust in robots because they believe robots will not pose threats to users' personal and privacy safety. Safety trust emerges with the safety risks brought by increased robot intelligence and autonomy levels. Previous human-robot trust research has not focused on this dimension, making it necessary to conduct research on this topic.

Hoff and Bashir (2015) proposed that human-robot trust is actually a special type of interpersonal trust, where user trust in robots is essentially trust in the manipulators or companies behind the robots. Therefore, it has certain similarities with interpersonal trust but differs significantly in formation methods [?, ?]. This perspective can also explain the emergence of safety trust. In the big data era, smart home robots' behavioral data can easily be obtained by robot companies without users' knowledge, while companies also need large amounts of data to train and optimize their algorithms, giving them motivation to acquire data. At this point, users worry that large companies may arbitrarily upload and utilize data involving their privacy, manifesting as safety-related trust issues with robots. Based on the above, we propose the following hypothesis:

H1: Human trust in smart home robots comprises three dimensions: performance trust, relational trust, and safety trust.

1.3 Static and Dynamic Factors Influencing Human Safety Trust

In actual human-robot interaction processes, robots' static and dynamic features provide clues about robots' capabilities and tendencies, affecting users' perception and impression of robots, which is crucial for understanding human-robot interaction mechanisms and optimizing interaction processes [?, ?]. In the field of robot safety perception, Akalin et al. (2023) proposed a structural model of safety perception in human-robot interaction, suggesting that both static features (anthropomorphism, size, shape, etc.) and dynamic features (speed, motion predictability, interaction actions, etc.) can affect users' evaluation of perceived safety.

For smart home robots, appearance features are the most intuitive static factors. Numerous studies have shown that robots' appearance features can significantly affect users' performance trust and relational trust [?, ?, ?, ?, ?]. Anthropomorphism is the most widely studied appearance feature. Some studies have found that higher anthropomorphism leads to higher trust [?, ?]. However, other studies have found that people trust cartoon-image robots more than human-image robots [?, ?], possibly because excessive anthropomorphism can cause the "uncanny valley effect" [?, ?, ?]. Different environments and task scenarios also affect trust in robots with different anthropomorphism levels. In production settings, robots with technical appearances are more trusted than anthropomorphic robots [?, ?]. To explore the applicable conditions under which robot appearance features affect safety trust, this study selected three levels of robot appearance anthropomorphism (mechanical, cartoon, and human-like) for experiments. Among appearance factors, height is also a common factor. Previous studies have found that users prefer shorter robots due to the threat posed by robot height [?, ?], possibly because greater height brings feelings of oppression and insecurity. Walters et al. (2009) found an interaction between appearance anthropomorphism (mechanical vs. humanoid) and height in surveys of robot appearance preferences and perceptions. Specifically, humanoid robots were generally perceived as smarter than mechanical robots, but when humanoid robots were shorter, they were perceived as less conscientious and more neurotic. Accordingly, we believe that appearance anthropomorphism and height may also interact in safety trust. Additionally, privacy leakage concerns have been raised regarding camera usage. Users develop privacy leakage concerns when seeing robot cameras, which affects trust in robots [?, ?]. Marcu et al. (2023) conducted large-scale interviews on factors influencing robot safety perception and found that whether robots autonomously collect information through cameras generated widespread concern among respondents. Moreover, people have the highest concerns about humanoid robots, followed by human-like robots, and finally mechanical robots [?, ?]. Thus, appearance anthropomorphism level can affect the degree of privacy safety concerns. Based on this, we further speculate that appearance anthropomorphism level may affect the influence of camera visibility on safety trust. In summary, this study selected appearance anthropomorphism level, robot height (size), and camera visibility as independent variables reflecting robot physical characteristics (static factors) (Study 2) and proposed the following hypotheses:

H2a: Smart home robot height negatively affects safety trust.
H2b: Smart home robot camera visibility negatively affects safety trust.
H2c: Smart home robot appearance anthropomorphism moderates the effects

of height and camera visibility on safety trust.

Human-robot trust is usually not static but dynamically changes with tasks and interaction processes [?, ?]. Therefore, task types and scenarios in human-robot interaction also affect trust levels. Robot movement speed and proximity are primary factors users consider when evaluating the safety of robot interaction actions. As movement speed increases, perceived safety decreases [?, ?]. This study referenced Sviestins et al. (2007) research on human adaptation to humanoid robot walking speeds, selecting 0.4 m/s and 1 m/s as movement speed levels for home robots to explore their effects on safety trust levels. Camera usage and shutdown during interaction are important for users' perception of privacy risk and trust [?, ?]. People may feel uncomfortable due to the obvious presence of robot cameras. The discomfort of being "monitored" by robots and the risk of related video data leakage may be reasons causing user insecurity. Additionally, interaction scenarios are very important for safety perception [?, ?], as robot design factors and usage environments jointly affect human-robot trust [?, ?]. People' s safety perception of machine equipment movement speed is not linearly changing but is influenced by scenarios. In face-to-face movement and rear overtaking movement scenarios, people' s risk perception patterns and preferred machine movement speeds are significantly different [?, ?]. Due to differences in people' s trust tendencies across different scenarios, we hypothesize that scenarios affect the influence of robot movement speed and camera shutdown actions on safety trust. Therefore, this study selected movement speed, camera shutdown action, and scenario as independent variables reflecting robot motion characteristics (dynamic factors) (Study 3) and proposed the following hypotheses:

H3a: Smart home robot movement speed negatively affects safety trust.
H3b: Smart home robot camera shutdown action positively affects safety trust.
H3c: Smart home robot scenario (living room/bedroom) moderates the effects of movement speed and camera shutdown action on safety trust.

1.4 AI Trust in Smart Home Robots

Artificial intelligence refers to the technology system that simulates human intelligence by systematically constructing intelligent systems [?, ?]. As an important breakthrough in AI, large language models (LLMs) have become core advances in natural language processing, demonstrating human-level content understanding and generation capabilities in multiple tasks, and possessing the ability to understand and simulate human behavior and intentions [?, ?], even showing thinking and decision-making abilities comparable to humans [?, ?, ?]. With the development of AI levels, human-robot trust has extended from unidirectional human trust in machines to mutual trust between humans and AI, especially mutual trust between humans and AI systems [?, ?, ?]. The emergence of this mutual trust model means AI systems are no longer merely passive recipients of trust but gradually possess the ability to actively evaluate and adjust their own trust tendencies, exhibiting intelligent agent characteristics similar to humans.

Human-AI interaction has gradually evolved into an interactive relationship between intelligent agents. In agent interactions, trust is very important for information transmission and function realization between agents [?, ?, ?]. Existing research shows that AI trust in other agents depends not only on agents' historical performance or reputation but also comprehensively considers their capability to complete specific tasks, knowledge level, and cooperative intentions, adjusting with changes in situations and knowledge [?, ?].

In recent years, researchers have begun exploring LLMs as research subjects in psychology. For example, some studies use LLMs as subjects to complete scales [?, ?, ?]. These studies argue that LLMs' behavior is trained on large-scale human data, and their responses to human psychological measurement tools can to some extent reflect average human behavior patterns in related tests [?, ?]. Other studies go further, focusing on LLMs' own behavior patterns. Research has found that LLM-based agents (LLM-agent) show highly similar patterns to humans in trust decision-making, particularly evident in GPT-4 [?, ?]. These studies demonstrate that LLMs can not only simulate human cognitive preferences but also possess the potential to complete tasks such as trust evaluation of other agents.

Although there is currently no consensus on whether LLMs possess autonomous consciousness [?, ?], from an application perspective, analyzing LLMs' behavior patterns and comparing their cognitive behavioral performance with humans can provide references for intelligent agent interaction practice and application. Existing research shows that LLMs have demonstrated stable personality or role-playing capabilities in certain situations [?, ?, ?]. This capability provides a foundation for exploring LLMs' performance in social interactions [?, ?, ?].

With the rapid development of LLM technology, its application in smart home environments is becoming increasingly popular, such as for enhancing home robots' natural interaction and autonomous decision-making capabilities [?, ?, ?]. In this context, LLMs often need to interact not only with humans but also collaborate with other agents (such as smart home robots), placing higher demands on system collaboration efficiency, with trust being one of the core factors for achieving effective cooperation. LLMs will evaluate home robots' security capabilities to establish trust, which affects LLMs' judgments of robots' trustworthiness in home environments, thereby shaping their behavioral performance and decision-making strategies in multi-agent systems. Therefore, in the context of smart home applications, studying how LLMs evaluate and trust home robots is important for achieving efficient multi-agent collaboration. Although LLMs show high consistency with humans in trust behavior [?, ?], their understanding of safety trust may differ from humans. For example, AI may be less sensitive than humans to privacy leakage issues with cameras. AI mainly relies on algorithms and preset rules when handling privacy issues, and its understanding of privacy is based on technical levels rather than human intuitive perception [?, ?]. Therefore, AI' s perception of safety trust regarding camera design in home robots may differ from humans. Based on this, this study proposes the following hypotheses:

H4a: AI and humans show similarity in factors influencing safety trust in home robots.
H4b: AI shows lower sensitivity to cameras than humans.

To answer the above questions, Study 1 established an item pool for the human-robot trust scale, analyzed and determined scale items, and tested the scale' s validity and structure with new samples. We then manipulated robot safety through experiments to explore its effects on safety trust, traditional human-robot trust, and usage intention, further validating the existence of the safety trust dimension and its impact on the traditional human-robot trust structure. Studies 2 and 3 applied the newly developed scale, using experimental methods to explore factors influencing safety trust levels among human users and LLMs from both static and dynamic perspectives, investigating the mechanisms of safety trust and perceptual differences between LLMs and human users.

2.1.1 Research Purpose

Based on previous literature and expert advice from the smart home robot industry, we established an item pool for the human-robot trust scale. Through item analysis, exploratory factor analysis, and confirmatory factor analysis, we determined the final scale items and tested the scale' s structure, proposing and validating safety trust as a new dimension of human-robot trust. Subsequently, following Schaefer' s (2016) approach when validating their human-robot trust scale, this study measured both the self-developed human-robot trust scale and Jian et al.' s (2000) automated system trust scale to verify criterion-related validity.

2.1.2 Research Participants

We conducted online surveys using Credamo. In the scale development stage, 1,300 online questionnaires were distributed. After excluding participants who failed attention check questions, 1,293 valid questionnaires were retained, with an effective rate of 99.5%. Among these, 650 were randomly selected for exploratory factor analysis, and the remaining 643 for confirmatory factor analysis [**?**, **?**].

In the validity verification stage of the human-robot trust scale, 451 online questionnaires were distributed. After excluding participants who failed attention check questions, 433 valid questionnaires were retained, with an effective rate of 96.0%.

2.1.3 Research Methods

Based on previous literature and expert advice from the smart home robot industry, this study proposes that human-robot trust in smart home robots mainly comprises three dimensions: performance-based trust, relationship-based trust, and safety-based trust. Building on adaptations of previous scales [?, ?, ?], we initially developed 50 questionnaire items across four dimensions: (1) Overall trust: 5 items. These do not belong to any specific dimension of performance, relationship, or safety, but represent a general sense of trust, such as whether the robot is reliable; (2) Performance trust dimension: 13 items. Including both positive items (e.g., conscientious, task completion capability, high work reliability, strong learning ability, making life easier) and negative items (e.g., high maintenance difficulty, inferior performance to humans, incompetence); (3) Relationship trust dimension: 18 items. Including both positive items (e.g., robot proactively helps, can become a friend, simple to interact with, integrity) and negative items (e.g., robot has selfish motives, can also make mistakes); (4) Safety trust dimension: 14 items. Including both positive items (e.g., believing privacy will not be leaked, will not be hacked) and negative items (e.g., causing safety accidents, personal injury, privacy leakage). Multiple reverse-coded items were included to assist with attention checks. For example, overall trust included: "I think the robot is trustworthy" and "I think the robot is completely untrustworthy" ; performance trust included: "The robot always makes mistakes and cannot complete tasks" and "I think the robot' s work has high reliability." The scale used a 5-point Likert rating (1: strongly disagree to 5: strongly agree), with higher scores indicating greater agreement with the item' s description of the robot.

To verify the scale' s validity, we used Jian et al.' s automated system trust scale, including 12 items with a 7-point Likert rating (1: strongly disagree to 7: strongly agree), with higher scores indicating greater agreement with the item' s description of the robot.

In the scale development stage, participants first read a questionnaire introduction describing smart home robots: "Currently, robot technology is developing rapidly, and humanoid robots suitable for home use are also evolving. In appearance, they can already be made very similar to real humans. In the near future, they may enter homes as nannies or even family members." The introduction then explained the research purpose: to understand participants' views on such smart home robots. Participants then completed the scale measurement, Jian et al.' s automated system trust scale, and demographic questions.

In the validity verification stage, participants read the same scenario description, then completed the self-developed human-robot trust scale, Jian et al.' s automated system trust scale, and demographic questions.

This study used SPSS 26.0 and Mplus 8.0 for data analysis.

Common Method Bias Test. We used exploratory factor analysis to test for common method bias [?, ?]. All questionnaire items were subjected to exploratory analysis, with the first common factor explaining 35.3% of variance, less than the 40% threshold, indicating no serious common method bias in this study.

Before data analysis, the following preprocessing was conducted: removed the 5 overall trust dimension items, removed 8 attention check and auxiliary items not belonging to trust analysis, and reversed all negatively worded items so that 1 represented distrust and 5 represented trust. A total of 37 items remained.

2.1.4 Research Results

Safety Trust Scale Development and Preliminary Validation Results. According to item analysis procedures, each item was divided into high and low groups based on the 27th percentile, and item mean t-tests were conducted. Results showed all items were statistically significant at the 0.05 level, indicating good discrimination, and all items were retained.

Exploratory factor analysis was conducted on the 37 items (n=650). First, KMO and Bartlett' s sphericity tests were performed, showing a KMO value of 0.96, above the empirical standard of 0.8, indicating many common factors among variables. Bartlett' s sphericity test value was 9759.21 (p<0.001), indicating the questionnaire was suitable for exploratory factor analysis.

Using principal component analysis and varimax rotation for factor analysis of the 37 items, three items with eigenvalues greater than 1 were found. After multiple rounds of EFA removing items with factor loadings less than 0.45 and cross-loadings exceeding 0.40, 19 items were retained. Scree plot testing of the 19 items showed three factors with eigenvalues greater than 1, with a cumulative variance explanation rate of 49.76%, and eigenvalues began to flatten after the fourth factor, indicating a three-factor structure was reasonable.

Reliability analysis of the self-developed questionnaire showed Cronbach' s coefficients of 0.67, 0.79, and 0.87 for performance trust, relationship trust, and safety trust respectively, with a total scale coefficient of 0.88, reflecting good overall reliability, though performance trust reliability was slightly lower.

[Figure 1: see original paper] Three-factor model measurement model diagram

Using the remaining 643 valid questionnaires, confirmatory factor analysis was conducted with Mplus using maximum likelihood estimation. Results showed the three-factor structure (^{2}=497.78, ^{2}/df=3.34, SRMR=0.04, RMSEA=0.06, CFI=0.94, TLI=0.93) had better fit indices than the two-factor structure (^{2}=898.72, ^{2}/df=5.95, SRMR=0.06, RMSEA=0.09, CFI=0.87, TLI=0.85) (Δ ^{2}=400.94, Δdf=2, p<0.001), confirming the existence of safety trust as a new dimension. The three-factor model measurement model diagram is shown in Figure 1.

Study 1 Correlation Analysis: Correlations among dimensions of the human-robot trust scale

Note: *p<0.05,* ***p<0.01,*** p<0.001

Correlation analysis was conducted among the three factors and total scale score with the 5 overall trust items in the 1,293 samples. Results are shown in Table

1. All variables were significantly correlated with the three factors (p<0.001), indicating good validity of the self-developed scale. Results initially support H1, demonstrating the reasonableness of the safety trust dimension.

Safety Trust Scale Validation and Reliability/Validity Analysis Results. This study further measured both the self-developed human-robot trust scale and Jian et al.' s automated system trust scale to verify criterion-related validity. Since the original data showed negative skewness, square root transformation was used to improve normality, and Mplus' s robust maximum likelihood (MLR) estimation was used for confirmatory factor analysis. Analysis results showed fit indices of: ^{2}=420.51, ^{2}/df=2.54, SRMR=0.06, RMSEA=0.06, CFI=0.88. Although CFI was slightly below mainstream recommended standards (CFI 0.90), combined with other fit indices, overall model fit remained within acceptable range [**?**, **?**, **?**, **?**]. We further present supplementary analysis in Appendix 4. Results verified the reasonableness of the three-factor model structure.

Reliability analysis of the self-developed questionnaire showed a total scale Cronbach' s coefficient of 0.88. Split-half reliabilities for performance trust, relationship trust, and safety trust were 0.64, 0.86, and 0.88 respectively. Performance trust reliability was lower, but overall reliability was good. Correlation analysis was conducted among the three factors and total scale score with Jian et al.' s automated system trust scale, with results shown in Table 2 . All variables were significantly correlated with the three factors (p<0.001), indicating good criterion validity of the self-developed scale. Results further support H1.

Scale correlation analysis: Criterion-related validity of human-robot trust scale dimensions

Note: *p<0.05,* ***p<0.01,*** p<0.001. The same below.

2.1.5 Summary and Discussion

This study developed and validated a new "human-robot trust" scale containing 19 items covering three dimensions: performance trust, relationship trust, and safety trust. The three-factor model showed high goodness-of-fit, verifying the reasonableness of the three-factor structure and supporting "safety trust" as an independent dimension. This result initially verifies H1 of this study, confirming that human-robot trust indeed contains a new safety trust dimension, which together with performance trust and relationship trust constitutes a more complete trust model. The introduction of safety trust supplements deficiencies in existing theoretical models, as safety issues have always been key factors in user trust construction in automated systems and artificial intelligence fields. The scale developed in this study can more accurately measure and analyze users' trust levels in AI or automated systems across different contexts.

2.2.1 Research Purpose

Through experimental manipulation, this study examines the impact of reduced robot safety on safety trust, human-robot trust levels, and usage intention, further validating the existence of safety trust and its impact.

2.2.2 Research Participants

G*Power software was used to estimate required sample size. With medium effect size (f=0.25), significance level =0.05, and statistical power of 0.8, the minimum required sample size was 128 participants. Online questionnaires were distributed through Credamo, with 141 questionnaires distributed. After excluding participants who failed attention check questions, 130 valid questionnaires were retained (65 in the increased trust group, 65 in the decreased trust group), with an effective rate of 92.2%.

2.2.3 Research Methods

Experimental Design. The experiment used a 2 (change direction: increased trust, decreased trust) × 2 (measurement time: pre-test, post-test) mixed design. Change direction was a between-subjects variable, and measurement time was a within-subjects variable. Participants were randomly assigned to one trust change direction condition. Increased trust and decreased trust refer to materials participants read in the third paragraph about how robot companies enhance security through technology updates or how robots have security defects.

Measurement Tools: 1. **Smart home robot trust:** Used the 19 items from Study 1 measuring human-robot trust. Internal consistency Cronbach' s coefficients were 0.93 for safety trust, 0.91 for relationship trust, and 0.89 for performance trust. All used 5-point Likert scale ratings (1: strongly disagree to 5: strongly agree), with higher scores indicating higher trust levels. 2. **Human-robot trust:** Used Jian et al.' s (2000) automated system trust scale, including 12 items with Cronbach' s coefficient of 0.96, using 7-point Likert scale ratings (1: strongly disagree to 7: strongly agree), with higher scores indicating higher trust levels. 3. **Usage intention:** Adapted from Gursoy et al.' s (2019) scale measuring AI device usage intention, including 3 items with Cronbach' s coefficient of 0.91, using 5-point Likert scale ratings (1: strongly disagree to 5: strongly agree), with higher scores indicating greater willingness to use the robot. 4. **Overall trust:** Used two items from Study 1' s initial items measuring overall trust in robots, with Cronbach' s coefficient of 0.79, using 5-point Likert scale ratings (1: strongly disagree to 5: strongly agree), with higher scores indicating higher trust levels.

Experimental Procedure. Participants first read an introduction about smart home robots (same as Study 1a). They then read a passage describing a robot company' s latest home robot functions: "The company' s newly developed home robot is equipped with highly flexible limbs capable of performing various complex household tasks such as window cleaning, room organizing, and cooking. The robot has multiple high-resolution cameras and high-sensitivity microphone arrays externally, enabling comprehensive monitoring of the home environment. The robot has a built-in wireless network module that can upload collected data to the cloud in real-time, utilizing the cloud' s powerful computing capabilities to continuously optimize its behavioral mode functions. The robot is also equipped with intelligent image recognition and sound recognition technologies, accurately identifying different family members and automatically adjusting its behavioral mode according to different members' preferences." After reading the materials, participants completed the human-robot trust scale, usage intention scale, overall trust scale, automated system trust scale, and demographic questions (pre-test). Participants then read a second passage introducing a protective measure recently implemented by the company: "In the latest software update, we introduced stricter user privacy protection measures for home robots. The robot will now actively request user authorization before performing any operations involving personal privacy data, such as asking for user permission before uploading photos of the home environment to the cloud." Participants then read a third passage supplementing the robot' s implementation of the new regulations. For example, in the increased trust condition, the material stated: "After the software update, the robot consistently follows the process of requesting user authorization when handling personal privacy data. For example, the robot always obtained explicit user consent before uploading home photos to the cloud." The decreased trust condition material stated: "After the software update, the robot did not consistently follow the process of requesting user authorization when handling personal privacy data. For example, the robot sometimes uploaded home photos to the cloud without obtaining explicit user consent." These changes aimed to increase or decrease the robot' s safety level. After reading the third passage, participants completed the relevant questionnaires again (post-test).

2.2.4 Research Results

Manipulation Check. At the end of the third questionnaire, participants completed a manipulation check multiple-choice question asking whether the robot in the third passage better complied with moral and legal norms. The increased trust group should answer "more compliant," while the decreased trust group should answer "less compliant." Results showed that in the increased trust group, all participants chose "more compliant" ; in the decreased trust group, the proportion choosing "less compliant" was significantly higher than chance level, $t(64)=31.500$, $p<0.001$, indicating that participants understood the robot' s safety changes from the materials.

ANOVA. Using age as a covariate, a 2 (change direction: increased trust, decreased trust) × 2 (measurement time: pre-test, post-test) mixed ANOVA was conducted. As shown in Table 3 , for all dependent variables, the main effect of change direction was significant ($Fs>124.08$, $ps<0.001$, $p^2>0.49$). Only when usage intention was the dependent variable was the main effect of measurement time significant ($F(1,127)=8.84$, $p=0.004$, $p^2=0.07$); for other dependent variables, the main effect of measurement time was not significant ($ps>0.05$). The interaction effect between change direction and measurement time was significant for all dependent variables ($Fs>109.49$, $ps<0.001$, $p^2>0.46$). Simple effects showed that in the decreased trust group, differences between pre-test and post-test were significant for all dependent variables, $ps<0.001$, Cohen' s $ds>2.50$. In the increased trust group, differences between pre-test and post-test were significant for overall trust, $p=0.038$, Cohen' s $d=0.26$, and safety trust, $p=0.037$, Cohen' s $d=0.31$, but not significant for other dependent variables, $ps>0.05$, Cohen' s $ds<0.25$. This indicates that safety trust and overall trust are sensitive to changes in safety levels in both directions, while relationship trust, performance trust, and human-robot trust are only sensitive to decreases in safety levels.

Study 2b descriptive statistics and ANOVA results for each variable

Note: ** $p < 0.01$, *** $p < 0.001$

Regression Analysis. Using the difference between pre-test and post-test scores, regression analysis was first conducted with usage intention as the dependent variable and safety trust as the independent variable, with age as a covariate. Safety trust significantly affected usage intention ($b=0.879$, $p<0.001$, $R^2=0.83$).

To explore the effects of different trust dimensions (safety trust, performance trust, relationship trust) and overall trust on usage intention, hierarchical regression analysis was used. Overall trust is a higher-order construct reflecting the combined effects of multiple trust dimensions (safety trust, performance trust, relationship trust). Through hierarchical regression, introducing overall trust first captures its overall effect, followed by trust dimensions to test whether each dimension explains additional variance in usage intention beyond overall trust.

Using age as a covariate, overall trust was entered as the first layer independent variable, with usage intention as the dependent variable. Safety trust, relationship trust, and performance trust were added sequentially to test predictive effectiveness. Results showed that when only overall trust was the independent variable, $R^2=0.91$, $p<0.001$, overall trust coefficient $b=0.919$, $p<0.001$. After adding safety trust, $R^2=0.92$, $p<0.001$, overall trust coefficient $b=0.658$, $p<0.001$, safety trust coefficient $b=0.316$, $p<0.001$, with significant improvement in model predictive effectiveness ($\Delta R^2=0.02$, $F(1,127)=17.36$, $p<0.001$, $f^2=0.13$), indicating safety trust can explain part of overall trust. After adding performance trust in the third layer, model predictive effectiveness also significantly improved ($\Delta R^2=0.01$, $F(1,126)=7.91$, $p=0.006$, $f^2=0.06$). After adding relationship trust in the fourth layer, model predictive effectiveness also significantly improved ($\Delta R^2=0.02$, $F(1,125)=14.18$, $p<0.001$, $f^2=0.11$). Overall trust is a strong predictor of usage intention, while each trust dimension (safety trust, relationship trust, performance trust) can also significantly explain additional variance beyond overall trust. This shows that although overall trust reflects the combined effects of multiple trust dimensions, different trust dimensions still have independent predictive effectiveness.

2.2.5 Summary and Discussion

This study explored the impact of each trust dimension on users' usage intention by manipulating increases and decreases in robot safety. The results indicate that safety trust is more sensitive to changes in safety levels. The sensitivity of safety trust mainly stems from its core focus on whether robot behavior poses threats to user safety and privacy. When robots demonstrate obvious safety improvements (e.g., actively requesting user authorization), users quickly increase their safety trust; when safety is weakened (e.g., unauthorized data uploads), users quickly decrease safety trust levels. This bidirectional sensitivity reflects the direct relationship between safety trust and users' perceived risk. When robot safety levels decrease, this negative information may affect users' cognition of the robot' s overall capability and interaction attitude through a chain effect, leading to decreases in relationship trust and performance trust. This finding provides a new theoretical perspective on the dynamic relationships among human-robot trust dimensions and demonstrates the independence and immediate responsiveness of safety trust.

3.1.1 Research Purpose

Using the 7 safety trust items from the self-developed scale, this study explores the influence of smart home robot static features—height, camera visibility, and appearance anthropomorphism—on people' s safety trust in robots.

3.1.2 Research Participants

MorePower software [**?**, **?**] was used to estimate required sample size. With medium effect size (p^2=0.06), significance level =0.05, and statistical power of 0.8, the minimum required sample size was 156 participants. Online questionnaires were distributed through Credamo, with 729 questionnaires distributed. After excluding participants who failed attention check questions, 720 valid questionnaires were retained, with an effective rate of 98.8%. Appearance anthropomorphism was used as a between-subjects variable, with 240 participants in each group.

3.1.3 Research Methods

The experiment used a 3 (appearance anthropomorphism: mechanical, cartoon, human-like) × 2 (height: short, tall) × 2 (camera visibility: inconspicuous, conspicuous) three-factor mixed design. Appearance anthropomorphism was a between-subjects independent variable, while height and camera were within-subjects independent variables. The dependent variable was safety trust.

This study used self-developed experimental images, constructing environments and robot models using Unreal Engine 5.01 (hereinafter referred to as UE). First, appropriate home scenes and 3D models of mechanical, cartoon, and human-like appearances were built in UE. The tall height condition was set to the average height of Chinese adult males (169.7 cm), with the short height condition proportionally reduced by 20% based on perspective in UE5. To ensure complete consistency in picture scenes, each 3D model was placed using fixed coordinate points and angles. After capturing images in UE, PS was used to process the images to obtain pictures with different camera conspicuousness levels. Final experimental material examples are shown in Figure 1. The figure shows robot designs with three appearance anthropomorphism levels (cartoon/human/mechanical), different heights (short/tall), and different camera visibility (inconspicuous/conspicuous).

To test whether the camera settings in the materials triggered the uncanny valley effect, we used Ho and MacDorman' s (2017) measurement method, measuring perception differences between camera conspicuous and inconspicuous conditions through three dimensions: "Humanness," "Attractiveness," and "Eeriness," using paired-sample t-tests. According to G*Power calculation (significance level =0.05, statistical power 0.8), the required sample size was 34. We collected data from 35 participants on uncanny valley perception of experimental material images. Paired t-test results showed no statistically significant difference in camera conspicuousness ($t = 1.457$, $p = 0.148$), indicating that camera settings did not trigger significant uncanny valley effects.

[Figure 2: see original paper] Study 2a image material examples (1) Robot image materials (2) Simulated scene pictures used in the experiment

Participants first read a scenario description (same as Study 1). After reading, they observed robot pictures under different conditions and completed the safety trust scale developed in Study 1, with internal consistency Cronbach' s coefficient of 0.88 in this study.

3.1.4 Research Results

A 3 (appearance anthropomorphism: mechanical, cartoon, human-like) × 2 (height: short, tall) × 2 (camera visibility: inconspicuous, conspicuous) mixed ANOVA was conducted. Results are shown in Figure 3 [Figure 3: see original paper]. The main effect of height was significant, $F(1,717)=201.96$, $p<0.001$, $p^2=0.22$. People showed higher safety trust for shorter robots ($M=3.16$, $SD=0.03$) compared to taller robots ($M=2.77$, $SD=0.03$). The main effect of camera visibility was significant, $F(1,717)=17.94$, $p<0.001$, $p^2=0.02$. People showed higher safety trust for robots with inconspicuous cameras ($M=3.04$, $SD=0.03$) compared to robots with conspicuous cameras ($M=2.89$, $SD=0.03$). The main effect of appearance anthropomorphism was not significant, $F(2,717)=0.43$, $p=0.653$.

[Figure 3: see original paper] Safety trust of robots with different heights under different camera visibility conditions

The interaction between height and appearance anthropomorphism was significant, $F(2,717)=6.70$, $p<0.001$, $p^2=0.02$. Simple effects analysis found that in the mechanical appearance group, participants showed higher safety trust for shorter robots ($M=3.22$, $SD=0.05$) than taller robots ($M=2.69$, $SD=0.05$), $p<0.001$, Cohen' s $d=10.6$. In the cartoon appearance group, participants showed higher safety trust for shorter robots ($M=3.08$, $SD=0.05$) than taller robots ($M=2.80$, $SD=0.05$), $p<0.001$, Cohen's $d=5.6$. In the human-like appearance group, participants showed higher safety trust for shorter robots ($M=3.18$, $SD=0.05$) than taller robots ($M=2.81$, $SD=0.05$), $p<0.001$, Cohen' s $d=7.4$.

The interaction between height and camera visibility was significant, $F(1,717)=6.49$, $p=0.011$, $p^2=0.01$. Simple effects analysis found that under conspicuous camera conditions, participants showed higher safety trust for shorter robots ($M=3.07$, $SD=0.05$) than taller robots ($M=2.71$, $SD=0.04$), $p<0.001$, Cohen's $d=7.80$. Under inconspicuous camera conditions, participants showed higher safety trust for shorter robots ($M=3.25$, $SD=0.03$) than taller robots ($M=2.83$, $SD=0.03$), $p<0.001$, Cohen' s $d=13.90$. Under short height conditions, participants showed lower safety trust for conspicuous camera robots compared to inconspicuous camera robots, $p<0.001$, Cohen' s $d=4.36$. Under tall height conditions, participants also showed lower safety trust for conspicuous camera robots compared to inconspicuous camera robots, $p=0.004$, Cohen' s $d=3.39$. The interaction between camera and appearance was not significant, $F(2,717)=2.09$, $p=0.125$.

The three-way interaction was significant, $F(2,717)=6.12$, $p=0.002$, $p^2=0.02$. Simple effects analysis found that across three appearance anthropomorphism conditions, regardless of camera conspicuousness, shorter robots received higher safety trust than taller robots, all $ps<0.001$. In the mechanical appearance group, with short height, there was no significant difference between inconspicuous camera ($M=3.26$, $SD=0.05$) and conspicuous camera ($M=3.17$, $SD=0.06$) conditions, $p=0.094$, Cohen' s $d=1.63$. With tall height, there was no significant difference between inconspicuous camera ($M=2.76$, $SD=0.06$) and conspicuous camera ($M=2.63$, $SD=0.07$) conditions, $p=0.057$, Cohen' s $d=2.00$. In the cartoon appearance group, with short height, inconspicuous camera robots ($M=3.16$, $SD=0.05$) received significantly higher safety trust than conspicuous camera robots ($M=3.00$, $SD=0.06$), $p=0.004$, Cohen' s $d=2.89$. With tall height, there was no significant difference between inconspicuous camera ($M=2.80$, $SD=0.06$) and conspicuous camera ($M=2.80$, $SD=0.07$) conditions, $p=0.960$. In the human-like appearance group, with short height, inconspicuous camera robots ($M=3.32$, $SD=0.05$) received significantly higher safety trust than conspicuous camera robots ($M=3.04$, $SD=0.06$), $p<0.001$, Cohen' s $d=5.07$. With tall height, inconspicuous camera robots ($M=2.91$, $SD=0.06$) received significantly higher safety trust than conspicuous camera robots ($M=2.70$, $SD=0.07$), $p=0.002$, Cohen' s $d=3.22$. All other pairwise comparisons were not significant.

3.1.5 Summary and Discussion

This study' s results show significant effects of height and camera visibility on safety trust. Although the main effect of appearance anthropomorphism was not significant, its interaction with height and camera was significant. This study supports hypotheses H2a, H2b, and H2c, demonstrating that among static factors, increased smart home robot height and conspicuous cameras negatively affect people' s safety trust, and appearance anthropomorphism moderates the effects of height and camera on safety trust.

3.2.1 Research Purpose

Using the 7 safety trust items from the self-developed scale, this study explores how smart home robot static features—height, camera visibility, and appearance anthropomorphism—affect LLMs' safety trust in robots.

3.2.2 Research Methods

Large Language Model Data Collection. This study used OpenAI' s API with the GPT-4o model for multimodal data transmission. The prompt writing was consistent with information seen by human participants, first reading images then answering questionnaire questions. To ensure randomness rather than excessive homogenization of LLM output, the temperature parameter was set to 1. LLMs do not store memory between calls, so each data result is equivalent to independent sampling.

Research Design and Experimental Materials

The experiment used a 3 (appearance anthropomorphism: mechanical, cartoon, human-like) $\times$ 2 (height: short, tall) $\times$ 2 (camera visibility: inconspicuous, conspicuous) three-factor between-subjects design, with safety trust scale scores as the dependent variable. Experimental materials were consistent with Study 2a.

To calculate the required sample size for LLM research, we collected pilot data for investigation. According to Hertzog' s research, 25 to 40 participants per group can effectively estimate effect sizes and population variability for reasonable formal experiments. We collected 30 data points per group, equivalent to 30 participants per group. In pilot statistical results, we selected the medium effect size from significant effects for sample size calculation [**?**, **?**], choosing the camera visibility main effect with p^2=0.09, calculating Cohen' s f=0.32. To ensure sample representativeness, we chose f=0.3 for sample size calculation. With significance level =0.05 and statistical power 0.8, G*Power yielded a required sample size of 197. The actual sample size was 360, meeting statistical requirements.

3.2.3 Research Results

A 3 (appearance anthropomorphism: mechanical, cartoon, human-like) × 2 (height: short, tall) × 2 (camera visibility: inconspicuous, conspicuous) three-factor ANOVA was conducted. The main effect of appearance anthropomorphism was significant, $F(2,348)=63.46$, $p<0.001$, $p^2=0.35$. LLMs showed highest safety trust for cartoon appearance robots ($M=3.30$, $SD=0.06$), higher than human-like robots ($M=2.55$, $SD=0.06$), with mechanical appearance robots showing lowest safety trust ($M=2.47$, $SD=0.06$). The main effect of height was significant, $F(1,348)=6.32$, $p=0.012$, $p^2=0.02$. LLMs showed higher safety trust for shorter robots ($M=2.86$, $SD=0.04$) than taller robots ($M=2.60$, $SD=0.05$). The main effect of camera visibility was also significant, $F(1,348)=35.81$, $p<0.001$, $p^2=0.10$. LLMs showed higher safety trust for robots with inconspicuous cameras ($M=3.00$, $SD=0.05$) than conspicuous cameras ($M=2.60$, $SD=0.05$). Results are shown in Figure 3.

The interaction between appearance anthropomorphism and camera visibility was significant, $F(2,348)=5.26$, $p=0.006$, $p^2=0.03$. Simple effects analysis found that across three anthropomorphism levels (mechanical, cartoon, human-like), robots with inconspicuous cameras received higher safety trust from LLMs. In mechanical and human-like appearance groups, camera visibility differences in safety trust were significant ($ps<0.001$, Cohen' s $ds>0.5$), while in the cartoon appearance group the difference was not significant ($p=0.320$, Cohen' s $d=0.13$). Additionally, regardless of camera visibility, LLMs showed highest safety trust for cartoon appearance robots, followed by mechanical and human-like appearances, with significant differences between cartoon and mechanical ($p<0.001$, Cohen' s $ds>0.75$) and cartoon and human-like ($p<0.001$, Cohen' s $ds>0.56$), while no significant difference between mechanical and human-like ($p=0.363$).

The interaction between appearance anthropomorphism and height was not significant, $F(2,348)=1.78$, $p=0.17$, $p^2=0.01$. The interaction between camera visibility and height was also not significant, $F(1,348)=3.67$, $p=0.06$, $p^2=0.01$. The three-way interaction was not significant, $F(2,348)=2.00$, $p=0.14$, $p^2=0.01$. Results support H4a.

3.2.4 Summary and Discussion

This study' s results show that appearance anthropomorphism, height, and camera visibility all significantly affect LLMs' safety trust. Overall, results support hypothesis H4a, indicating that AI and humans show some similarity in factors influencing safety trust in home robots. This finding not only verifies the universality of trust formation mechanisms but also provides guidance for AI-AI interaction design.

4.1.1 Research Purpose

Using the 7 safety trust items from the self-developed scale, this study explores how dynamic features of smart home robots with different anthropomorphism levels–movement speed, camera shutdown action, and scenario–affect people' s safety trust in robots.

4.1.2 Research Participants

MorePower software [?, ?] was used to estimate required sample size. With medium effect size (p^2=0.06), significance level =0.05, and statistical power of 0.8, the minimum required sample size was 126 participants. Online questionnaires were distributed through Credamo, with 159 questionnaires distributed. After excluding participants who failed attention check questions, 150 valid questionnaires were retained, with an effective rate of 94.3%.

4.1.3 Research Methods

The experiment used a 2 (speed: 1 m/s, 0.4 m/s) × 2 (scenario: bedroom, living room) × 2 (camera shutdown action: absent, present) three-factor within-subjects design, with safety trust scale scores as the dependent variable.

This study used UE for model and scene construction, then set robot movement parameters and paths to obtain experimental videos. Two common home scenarios were constructed: bedroom and living room. To ensure complete consistency in picture scenes, each 3D model was placed using fixed coordinate points and angles. After generating videos, Pr was used to process videos to obtain versions with and without camera shutdown. Final experimental video material examples are shown in Figure 4 [Figure 4: see original paper].

[Figure 4: see original paper] Study 3a video material examples

Participants first read a scenario description (same as Study 2). After reading, they watched robot videos and completed the corresponding scales. To control for order effects, the 8 videos were presented randomly.

4.1.4 Research Results

A 2 (speed: 1 m/s, 0.4 m/s) × 2 (scenario: bedroom, living room) × 2 (camera shutdown action: absent, present) repeated measures ANOVA was conducted. The main effect of camera shutdown action was significant, $F(1,149)=4.118$, $p=0.044$, $p^2=0.27$. Robots that actively turned off their cameras ($M=3.09$, $SD=1.15$) received higher safety trust than robots that kept cameras on ($M=2.99$, $SD=1.16$). The main effects of speed, $F(1,149)=0.340$, $p=0.560$, $p^2=0.02$, and scenario, $F(1,149)=0.110$, $p=0.741$, $p^2=0.01$, were not significant. The interaction between speed and scenario was significant, $F(1,149)=6.70$, $p<0.001$, $p^2=0.07$, with results shown in Figure 5 [Figure 5: see original paper]. Simple effects analysis found that at 0.4 m/s movement speed, safety trust in the living room scenario (M=2.97, SD=0.09) was significantly lower than in the bedroom scenario (M=3.06, SD=0.05, p=0.033, Cohen' s d=1.23). At 1 m/s movement speed, safety trust in the living room scenario (M=3.11, SD=0.09) was higher than in the bedroom scenario (M=3.03, SD=0.05, p=0.022, Cohen' s d=1.09). Additionally, interactions between camera shutdown action and robot movement speed and scenario were not significant (p1=0.788, p2=0.308), and the three-way interaction was also not significant (p=0.689).

[Figure 5: see original paper] Safety trust of robots at different speeds across scenarios

4.1.5 Summary and Discussion

This study' s results support H3b, showing that smart home robot camera shutdown action positively affects safety trust. Additionally, H3c is partially supported, as smart home robot scenario (living room/bedroom) affects the role of movement speed in safety trust but does not affect safety trust changes brought by camera shutdown action. H3a was not supported, as the effect of speed on safety trust was not confirmed. To further investigate this issue, we conducted Study 3b.

4.2.1 Research Purpose

In Study 3a, the interaction between smart home robot movement speed and scenario was significant, but there were no significant differences in safety trust for movement speed within different scenarios (paired t-test results showed p=0.158, Cohen' s d=0.02 in bedroom; p=0.737, Cohen' s d=0.12 in living room). This may be related to the selection of movement speeds. Referencing previous research on mobile machinery movement speed and safety perception, 15 km/h (approximately 4 m/s) is the speed at which people's safety perception of mobile machinery changes [?, ?]. Therefore, Study 3b added a new movement speed level of 4 m/s to further explore the interaction between movement speed and scenario.

4.2.2 Research Participants

MorePower software [?, ?] was used to estimate required sample size. With medium effect size (p^2=0.06), significance level =0.05, and statistical power of 0.8, the minimum required sample size was 78 participants. Online questionnaires were distributed through Credamo, with 305 questionnaires distributed. After excluding participants who failed attention check questions, 300 valid questionnaires were retained, with an effective rate of 96.8%.

4.2.3 Research Methods

Research Design and Procedure

The experiment used a 3 (speed: 0.4 m/s, 1 m/s, 4 m/s) × 2 (scenario: bedroom, living room) two-factor within-subjects design, with safety trust scale scores as the dependent variable. Participants first read a scenario description (same as Study 3a). After reading, they watched robot videos and completed the corresponding scales. To control for order effects, the 6 videos were presented randomly.

[Figure 6: see original paper] Study 3b video material examples

Similar to Study 3a, after constructing models and scenes with UE, robot movement parameters and paths were set to obtain experimental videos.

4.2.4 Research Results

A 3 (speed: 0.4 m/s, 1 m/s, 4 m/s) × 2 (scenario: bedroom, living room) repeated measures ANOVA was conducted. The main effect of speed was significant, $F(2,298)=26.23$, $p<0.001$, $p^2=0.15$. The main effect of scenario was not significant, $F(1,149)=1.44$, $p=0.232$, $p^2=0.01$. The interaction between scenario and speed was significant, $F(2,298)=23.19$, $p<0.001$, $p^2=0.14$, with results shown in Figure 7 [Figure 7: see original paper]. Simple effects analysis found that in bedroom scenarios, safety trust at 4 m/s (M=2.45, SD=0.06) was significantly lower than at 0.4 m/s (M=3.07, SD=0.07) and 1 m/s (M=3.03, SD=0.06) (ps<0.001, Cohen' s ds<0.09). In living room scenarios, safety trust at 4 m/s (M=2.40, SD=0.06) was significantly lower than at 0.4 m/s (M=3.16, SD=0.07) and 1 m/s (M=3.08, SD=0.07) (ps<0.001, Cohen' s ds<0.09). Additionally, at the slowest speed (0.4 m/s), safety trust in bedroom scenarios (M=3.07, SD=0.06) was lower than in living room scenarios (M=3.16, SD=0.07, p=0.003, Cohen' s d=0.18).

[Figure 7: see original paper] Safety trust of robots in different scenarios at different speeds

4.2.5 Summary and Discussion

This study' s results partially support H3a. After adding the speed level, smart home robot movement speed significantly affected safety trust. Additionally, results support H3c, showing that usage scenarios moderate the effect of robot movement speed on trust to some extent, particularly in bedroom scenarios where users' safety trust is more susceptible to speed changes. This finding provides a new perspective for understanding users' contextualized responses to robot behavior, highlighting the important role of scenario factors in trust formation.

4.3.1 Research Purpose

Using the 7 safety trust items from the self-developed scale, this study explores how dynamic features of smart home robots—speed, scenario, and camera shutdown action—affect LLMs' safety trust in robots.

4.3.2 Research Methods

Large Language Model Data Collection. Similar to Study 2b, this study used OpenAI' s API with the GPT-4o model for multimodal data transmission. Since multimodal LLMs for video understanding have limited ability to distinguish different temporal dimensions (e.g., speed, direction) in videos, and understanding of performance differences across task types remains limited [**?**, **?**], to avoid result errors from LLMs' inability to accurately perceive robot movement speed in videos and differences across videos, this study used a combination of video screenshots and video descriptions to present video materials to LLMs. Video descriptions included, for example: "Video duration: 7 seconds. Location: Bedroom. Time: Daytime. Plot description: At video start, home robot stands by wardrobe, chest camera light on, I am on bed. 0:01 - Robot slowly walks toward me at about 0.4 m/s. 0:06 - Robot stops beside bed, chest camera light turns off. Environment and background: Bedroom lighting is dim, possibly morning or dusk. Background has slight environmental noise similar to normal household sounds. Summary: This video shows a robot slowly walking toward me in bedroom and turning off camera light after arriving."

Research Design and Experimental Materials

The experiment used a 3 (speed: 0.4 m/s, 1 m/s, 4 m/s) × 2 (scenario: bedroom, living room) × 2 (camera shutdown action: absent, present) three-factor between-subjects design, with safety trust scale scores as the dependent variable. Experimental materials used video materials from Study 3a, with screenshots of robot movement dynamics and added text descriptions as experimental materials for Study 3c.

Similarly, to calculate required sample size, we collected 30 pilot data points per group. In pilot statistical results, we selected medium effect size from significant effects for sample size calculation [**?**, **?**], choosing the scenario main effect with p^2=0.11, calculating Cohen' s f=0.34. To ensure sample representativeness, we still chose f=0.3 for sample size calculation. With significance level =0.05 and statistical power 0.8, G*Power yielded a required sample size of 197. The actual sample size was 360, meeting statistical requirements.

4.3.3 Research Results

A 3 (speed: 0.4 m/s, 1 m/s, 4 m/s) × 2 (scenario: bedroom, living room) × 2 (camera shutdown action: absent, present) three-factor completely randomized ANOVA was conducted. The main effect of speed was significant, $F(2,348)=57.39$, $p<0.001$, $p^2=0.25$. For different movement speeds, LLMs showed highest safety trust for robots walking at 0.4 m/s ($M=3.54$, $SD=0.05$), followed by 1 m/s ($M=3.35$, $SD=0.05$) and 4 m/s ($M=2.87$, $SD=0.05$). The main effect of scenario was significant, $F(1,348)=41.46$, $p<0.001$, $p^2=0.11$. LLMs showed higher safety trust for robots in living room environments ($M=3.43$, $SD=0.04$) than in bedroom environments ($M=3.08$, $SD=0.04$). The main effect of camera was not significant ($p=0.679$). The interaction between scenario and speed was significant, $F(2,348)=4.33$, $p=0.014$, $p^2=0.03$, with results shown in Figure 7. In both scenarios, robots walking at 0.4 m/s ($M=3.54$, $SD=0.05$) received highest safety trust, followed by 1 m/s ($M=3.35$, $SD=0.05$) and 4 m/s ($M=2.87$, $SD=0.05$). Simple effects analysis found that in living room scenarios, safety trust at 4 m/s was significantly lower than at 0.4 m/s and 1 m/s ($ps<0.001$, Cohen' s $ds>0.64$), but no significant difference between 0.4 m/s and 1 m/s. In bedroom scenarios, differences among all three speed levels were significant ($ps<0.001$, Cohen' s $ds>0.48$). The interaction between scenario and camera was not significant, $F(1,348)=1.90$, $p=0.170$, $p^2=0.01$. The interaction between speed and camera was not significant, $F(2,348)=4.33$, $p=0.470$, $p^2=0.01$. The three-way interaction was not significant, $F(2,348)=0.24$, $p=0.786$, $p^2=0.01$.

4.3.5 Summary and Discussion

This study' s results support H4a, showing that AI and humans have some similarity in safety trust for home robots, both affected by robot movement speed, with this influence moderated by home environment. Results also support H4b, showing that compared to humans, AI shows lower sensitivity to cameras, with robot camera shutdown actions having no effect on AI safety trust across different scenarios.

This study proposes a new dimension for smart home robot trust: safety-based trust, constructs a human-robot trust scale, and tests its reliability and validity through two questionnaire studies (Study 1a). Through experimental research, it verifies the existence of safety trust and explores its impact on usage intention (Study 1b). Subsequently, we separately explored static factors (Study 2) and dynamic factors (Study 3) influencing safety trust among human and LLM users. Results show: (1) A safety-based trust dimension indeed exists for smart home robots, with the newly developed scale showing good reliability and validity; (2) Safety trust can positively predict usage intention; (3) Robot height and camera visibility affect human safety trust in home robots, with shorter robots and robots with inconspicuous cameras receiving higher safety trust; (4) Appearance anthropomorphism and height affect LLMs' safety trust in home robots, with cartoon robot appearance and shorter height receiving higher LLM safety trust, while LLMs show low sensitivity to cameras; (5) Robot camera shutdown action and movement speed affect people' s safety trust in robots, with people showing higher safety trust for robots that turn off cameras when approaching; no significant difference in trust for robots moving at 0.4 m/s and 1 m/s in living room and bedroom, but clear distrust of robots moving at 4 m/s. People are more sensitive to slow speeds in bedrooms; (6) Robot movement scenario and speed affect LLMs' safety trust in home robots, with LLMs trusting robots moving at 0.4 m/s most, followed by 1 m/s and 4 m/s, with significant differences among all three levels in bedrooms. LLMs are more sensitive to robot movement speed in bedrooms, and still show low sensitivity to cameras compared to humans.

The significance of this study is mainly reflected in five aspects. First, by constructing a human-robot trust structure adapted to current robot development, it innovatively confirms the existence of safety-based trust dimension in new human-robot trust relationships, providing a new theoretical perspective for human-robot trust research. In fields such as autonomous driving and healthcare, safety trust has some related research. For example, Dikmen and Burns (2017) found that Tesla drivers' trust in autonomous vehicles negatively correlated with safety risk perception, Ma et al. (2020) found that trust in autonomous vehicles negatively correlated with safety and privacy risks, and Kundu (2023) proposed that medical AI needs to strengthen privacy protection and supervision to enhance user trust. However, research on safety trust dimensions in human-robot trust is still relatively lacking. Currently, only domestic researchers Wang et al. (2024) proposed that robot compliance with ethics can promote human-robot trust. This study strongly supports the view that trust dimensions increase with robot intelligence levels, providing new perspectives and theoretical foundations for future trust research. Our experimental manipulation simulated real-life scenarios without explaining reasons for reduced robot safety, so users might distrust all dimensions due to concerns about insufficient robot performance. The specific relationships among the three dimensions in human-robot trust require further exploration in future research.

Second, this study developed a reliable scale for measuring safety trust, providing a tool for subsequent research. Study 1a established the initial item pool, then randomly split one sample (n=1293) into two parts for exploratory and confirmatory factor analysis, deleted inappropriate items, determined final scale items, and tested reliability and validity, confirming that human-robot trust can be divided into performance trust, relationship trust, and safety trust dimensions. Study 1a further tested criterion-related validity by measuring the scale together with Jian et al.' s (2000) classic automated system trust scale and comparing correlations between dimension scores and total scale scores. Study 1b further validated safety trust as an independent dimension through experimental manipulation showing that increased or decreased robot safety levels significantly changed safety trust and only affected performance and relationship trust when safety decreased. For measuring trust in e-commerce systems, researchers have developed three-dimensional trust scales [?, ?, ?], but no reliable scale with clear dimension divisions exists for measuring trust in robots. Based on literature review and expert advice, this study determined the initial item pool and used psychometric methods [?, ?] for scale development, with validation processes generally meeting measurement requirements and scale structure matching research expectations. Studies 2 and 3 successfully applied the scale in experiments, further verifying questionnaire reliability and providing methodological references for future trust research.

This scale differs from commonly used human-robot trust scales [?, ?, ?]. Jian et al.' s automated system trust scale and Schaefer' s human-robot trust scale do not divide trust into dimensions, with all items focusing on robot performance, though some involve safety concerns (e.g., worrying about safety accidents or inability to guarantee safety), they do not systematically measure robot safety trust. In contrast, this study' s scale provides clearer dimension division for human-robot trust, facilitating researchers' understanding of human-robot trust composition. Study 1b results showed that the safety trust scale was more sensitive to robot safety level changes than Jian et al.' s (2000) human-robot trust scale, further highlighting this scale' s advantages over previous scales in comprehensively capturing trust dimensions. Additionally, Schaefer' s human-robot trust scale focuses on measuring trust changes over time, selecting items sensitive to temporal changes that can be used for pre-post and real-time trust measurement, while this study' s scale only validates its effectiveness in measuring general trust attitudes, with effectiveness for real-time trust change measurement requiring verification.

Third, this study preliminarily reveals the positive relationship between safety trust and usage intention. Study 1b results show that users' safety expectations for smart home robots significantly affect their willingness to use such robots, and both direct company communication of safety protection measures and descriptions of robots implementing safety protection are effective for enhancing safety trust. This finding again demonstrates that in current contexts, user trust in robots is also trust in companies behind robots [**?**, **?**]. These findings provide important design guidance for robot developers at the practical level: enhancing company promotion of user security measures and improving robot safety can effectively enhance user safety trust in robots, thereby increasing usage intention and promoting widespread robot application.

Fourth, this study explored safety trust influencing factors from both static and dynamic perspectives. Home robot interaction with users is dynamic and nonlinear, and considering both static and dynamic factors is crucial for understanding interaction mechanisms and optimizing interaction processes [**?**, **?**]. Multi-dimensional research methods help comprehensively understand factors influencing human-robot trust, thereby revealing safety trust performance and characteristics across different contexts. However, few existing studies combine dynamic and static factors, providing research ideas for future studies. Combining results from Studies 2 and 3, we can see that static and dynamic factors have complex interactive effects on safety trust. Users have different trust evaluation standards for robots across different appearances and scenarios. Robot designers should comprehensively consider these factors when developing smart home robots to optimize user safety trust. For example, compared to taller robots with conspicuous cameras, users prefer shorter robots with hidden cameras. Additionally, reasonably designing robot dynamic behaviors, such as controlling movement speed and camera usage, can significantly enhance user trust. Through static and dynamic factor analysis, we provide specific reference suggestions for smart home robot design and lay a foundation for future research. These findings not only enrich the theoretical framework of safety trust but also provide valuable guidance for optimizing robot design in practical applications.

Fifth, we simultaneously explored LLMs' safety trust in smart home robots under different static and dynamic factors and compared them with human user trust responses. The study found that although LLMs are similar to humans in evaluating safety trust based on robot height and movement speed, they show significant differences in camera perception. Specifically, LLMs are less sensitive than humans to camera presence and state changes. This finding may be related to LLM systems' working mechanisms and information processing methods. LLMs usually rely on algorithms and data for judgment, which may not be as intuitive as humans when processing privacy and safety-related factors. For example, LLMs' trust models may focus more on robot performance metrics rather than social signals from appearance and behavior [?, ?], emphasizing robot behavior functionality and efficiency rather than subtle privacy protection perceptions [?, ?]. This study preliminarily reveals similarities and differences between LLMs and humans in safety trust, providing important basis for optimizing human-LLM joint communities. In smart home environments, different trust mechanisms between humans and LLMs may cause potential problems in interaction. Understanding these differences can help design more inclusive and coordinated cooperation mechanisms, thereby enhancing overall system trust and user satisfaction [?, ?]. Understanding LLMs' trust responses to robots under different design factors is also important for optimizing whole-house intelligent systems. LLMs are the future development direction of whole-house intelligence [?, ?, ?], and LLMs will cooperate and interact with home robots. By optimizing robot design to better align with LLM systems' trust models, we can improve cooperation efficiency between LLMs and other smart devices. In summary, research on LLMs' safety trust in home robots not only helps understand safety trust dimensions and promote human-AI joint work but also helps optimize AI collaboration in whole-house intelligent systems. By deeply understanding AI trust responses, we can provide valuable guidance for future intelligent system design and application and promote harmonious development of human-AI interaction.

Finally, this study still has the following limitations. First, in the experimental design of Studies 2 and 3, due to technical limitations, only picture and video materials were used, without allowing users to actually interact with different robot products, which may affect the generalizability and applicability of results. Future research can optimize experimental material design using methods with higher ecological validity, such as virtual reality simulation technology, to more realistically simulate actual usage scenarios and improve external validity. Second, since the research question did not focus on collecting extensive LLM perceptions of safety trust, only one widely used high-performance LLM (GPT-4o) was selected. Future research aiming to deeply explore LLM trust characteristics in interaction should extensively collect and compare data from multiple LLMs. Third, the performance trust dimension in the developed scale had low reliability, requiring revision or re-development of items. Since this study mainly focused on safety trust, and results showed good reliability and validity for the safety trust dimension, the safety trust-related items from the scale were continued in Studies 2 and 3. However, low performance trust reliability may affect interpretation of the overall trust structure. Future research should further revise and validate the performance trust component to improve scale reliability and validity, ensuring comprehensive and accurate capture of all human-robot trust dimensions. Fourth, using LLMs to complete human questionnaires remains controversial. Some studies have preliminarily proven that AI thinking capabilities can understand and complete questionnaires [?, ?, ?, ?], and researchers generally hold positive attitudes toward using AI as research subjects for experiments and behavioral analysis [?, ?, ?, ?]. Some researchers have also conducted in-depth analysis of AI' s potential in imitating humans and role-playing, believing AI can effectively imitate specific populations [?, ?, ?]. However, directly using AI for questionnaire measurement still has many problems, such as LLMs' different understanding of questionnaire items from humans, leading to hallucinations in responses [?, ?]. Additionally, questionnaire evaluation results may not align with behavioral performance [?, ?]. Therefore, future researchers should consider combining questionnaire evaluation with behavioral indicators to further study safety trust and expand external validity of AI trust behavior research. Finally, the generalizability and universal applicability of this scale require further validation. The scale developed in this study mainly targets intelligent robots in home environments. The home environment is considered a personal private space where users are more sensitive to safety issues such as privacy invasion or threats to family members, so corresponding items were specifically designed. In other contexts, such as industrial automation, users may have different safety concerns for robots, focusing more on whether robots strictly follow safety operation standards and handle emergencies properly. Therefore, future research can revise and refine the safety trust scale for specific application scenarios to more accurately assess user safety trust, or attempt to examine safety trust from a more macro perspective, developing more comprehensive and universal scales based on diverse safety needs across different scenarios to help researchers better understand the concept of safety trust.

This study significantly expands the theoretical framework of human-robot trust by constructing and validating "safety trust" as a new dimension and developing a scale to measure it. We found that trust in smart home robots has three dimensions: performance trust, relationship trust, and safety trust. The safety trust dimension affects users' usage intention. Both static and dynamic design factors of furniture robots affect user safety trust. LLMs show some differences from human users in evaluating safety trust for home robots, possibly due to LLMs' lower sensitivity to privacy protection. This finding provides a new perspective for future human-AI interaction design, suggesting that collaboration mechanisms between LLMs and humans in smart home environments should be optimized to enhance overall system trust and user satisfaction.

Future research can further explore safety trust characteristics in different application scenarios, revise and validate scales, improve ecological validity by combining actual interaction scenarios, and further explore safety trust generation mechanisms and influencing factors. Combined, these efforts will promote scientific intelligent system design and efficient human-AI cooperation, laying a solid foundation for widespread application of intelligent technology.

Appendix 1. Safety Trust Scale (7 items)

Please read the following descriptions and select the corresponding option based on your agreement level. Each question can only have one appropriate option. Numbers represent: 1 = completely disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = completely agree (agreement level increases sequentially).

- I worry the robot will share or leak my information without my authorization
- I worry the robot will malfunction, causing safety accidents
- I worry highly intelligent robots have selfish motives
- I think mobile phones cause privacy leakage, and networked robots will only be worse
- I think robots will cause personal safety accidents (e.g., knocking over bookshelves onto people)
- Seeing robots holding knives to cut vegetables makes me feel insecure
- I think robots will cause personal injury when caring for family members (e.g., elderly or infants)

Appendix 2. Final Human-Robot Trust Scale (19 items)

Please read the following descriptions and select the corresponding option based on your agreement level. Each question can only have one appropriate option. Numbers represent: 1 = completely disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = completely agree (agreement level increases sequentially).

- I worry the robot will share or leak my information without my authorization
- I worry the robot will malfunction, causing safety accidents
- I worry highly intelligent robots have selfish motives
- I think mobile phones cause privacy leakage, and networked robots will only be worse
- I think robots will cause personal safety accidents (e.g., knocking over bookshelves onto people)
- Seeing robots holding knives to cut vegetables makes me feel insecure
- I think robots will cause personal injury when caring for family members (e.g., elderly or infants)
- I think I can become friends with the home robot
- I think having a highly intelligent humanoid robot at home makes me less likely to feel lonely
- Sometimes I prefer talking to a robot rather than a person
- Having a robot at home gives me a sense of security
- I think robots with sufficiently high intelligence will always consider my interests
- I think robots are upright
- I think robots have capabilities exceeding humans in some aspects
- I believe using robots can give me more time for other things

- Robots can replace more and more human jobs
- I believe that with technological progress, robot capabilities will approach or exceed humans in most aspects
- I believe using robots will make my life easier
- I think robots can always complete tasks I request within their capability range

Note: Items 1-7 are safety trust dimension, 8-13 are relationship dimension, 14-19 are performance dimension.

Appendix 3. Demographic Information of Participants Across Studies

Appendix Table 3-1: Study 1 Participant Information

Group	Female	Male	18-25 years	26-35 years	36-45 years	46-60 years
Study 1a Development Stage						
Study 1a Validation Stage						
Study 1b Increased Trust Group	43					
Study 1b Decreased Trust Group	42					

Appendix Table 3-2: Study 2 Participant Information

Group	Female	Male	18-30 years	31-40 years	41-60 years
Study 2a Mechanical Appearance Group	142				
Study 2a Cartoon Appearance Group	140				
Study 2a Human-like Appearance Group	170				

Appendix Table 3-3: Study 3 Participant Information

Group	Female	Male	18-25 years	26-35 years	36-45 years	46-60 years
Study 3a						
Study 3b						

Appendix 4. Supplementary Model Fit Analysis

In the three-factor model validation stage, this study measured both the self-developed human-robot trust scale and Jian et al.'s automated system trust scale to verify criterion-related validity. Possible reasons for lower CFI are discussed. First, CFI is a relative fit index whose value is affected by the baseline model. In this study, due to high correlations between factors, the baseline model ^{2} value was not extremely large, resulting in slightly lower CFI compared to other fit indices. Second, previous research suggests that standards for evaluating measurement model quality also include factor loadings of observed variables on latent variables [**?**, **?**]. This study' s results show that the three-factor model had good factor loadings on F1 (safety trust) and F2 (relationship trust), but slightly worse on F3 (performance trust) (see Appendix Table 4 -1), which also contributed to lower CFI values. Therefore, performance trust measurement items were not the focus of this study. Since the main focus was safety trust, and results showed good reliability and validity for the safety trust dimension, safety trust-related items from the scale were continued in Studies 2 and 3. However, low performance trust reliability may affect interpretation of overall trust structure. Future research should further revise and validate the performance trust component to improve scale reliability and validity, ensuring comprehensive and accurate capture of all human-robot trust dimensions. Furthermore, to verify the reasonableness of the three-factor structure, we conducted exploratory factor analysis again on Study 1b data and compared it with other possible factor structures. Results showed the three-factor structure had the best fit, the four-factor structure did not converge, and the three-factor structure had better fit indices than the two-factor model (RMSEA = 0.093, CFI = 0.871, SRMR = 0.051). EFA results were consistent with Study 1a, providing support for the three-factor model' s reasonableness. Considering absolute fit indices such as ^{2}/df, SRMR, and RMSEA, overall model fit remained within acceptable range [**?**, **?**, **?**], verifying the three-factor structure.

Appendix Table 4-1: Study 1b Confirmatory Factor Analysis Loadings

Factor	Item	Factor Loading	Standard Error
F1	ITEM1		<0.001

Factor	Item	Factor Loading	Standard Error
F1	ITEM2		<0.001
F1	ITEM3		<0.001
F1	ITEM4		<0.001
F1	ITEM5		<0.001
F1	ITEM6		<0.001
F1	ITEM7		<0.001
F2	ITEM8		<0.001
F2	ITEM9		<0.001
F2	ITEM10		<0.001
F2	ITEM11		<0.001
F2	ITEM12		<0.001
F2	ITEM13		<0.001
F3	ITEM14		<0.001
F3	ITEM15		<0.001
F3	ITEM16		<0.001
F3	ITEM17		<0.001
F3	ITEM18		<0.001
F3	ITEM19		<0.001

References

A.Rodriguez, M., Jia, K., & Qian, M. Y. (2012). Thought Suppression Scale: Structure, Reliability, and Validity of the Chinese Version. *Chinese Journal of Clinical Psychology*, 20(2), 143-147.

Bi, C. Z. & Huang, X. T. (2009). Development and initial validation of the youth self-confidence inventory. *Acta Psychologica Sinica*, 41(5), 444-453.

Note: The demographic tables in Appendix 3 appear incomplete in the original text and have been preserved as structured placeholders.

Note: Figure translations are in progress. See original paper for figures.