International Journal of Social Robotics
https://doi.org/10.1007/s12369-022-00961-3

The Hitchhiker's Guide to a Credible and Socially Present Robot: Two Meta-Analyses of the Power of Social Cues in Human–Robot Interaction

Kun Xu[1] · Mo Chen[1] · Leping You[2]

Accepted: 22 December 2022

Abstract

Social cues have been construed as an important concept in human–robot interaction, as they can be manipulated to reflect robots' perceived genders, personalities, emotions, identities, and so on. This study seeks to understand the overall effects of social cues and applies two meta-analyses to explore a hierarchy of social cues that elicits different degrees of users' social responses. A total of 25 and 44 effect sizes were calculated to represent the respective magnitudes of the effects of social cues on users' social presence ($N = 2498$) and trust in social robots ($N = 4147$). Results suggested that although the overall effects of social cues were small, manipulating social robots' facial and kinetic cues can induce medium-to-large-sized effects on users' social presence and trust. In addition, the overall positive effect sizes of social cues indicated that designing humanlike, natural, and lifelike cues was effective in evoking users' social presence and trust in social robots. The results of the two meta-analyses can contribute to the theoretical implications of the Computers are Social Actors paradigm and the practical and methodological design of human–robot interaction.

Keywords Social cues · Social robots · Social presence · Trust · Meta-analysis · Computers are Social Actors paradigm · Artificial intelligence

1 Introduction

The Computers are Social Actors (CASA) paradigm was proposed in the early 1990s to describe a series of human—computer interaction (HCI) practices in which users treat computers as if they were social actors. As some examples, past work has suggested that users apply politeness rules to interaction with computers [91], assign gender stereotypes to computers [89], and demonstrate strong preferences for collaboration with computers that are perceived as teammates [90]. Based on the findings of the CASA paradigm, Reeves and Nass expanded the research scope to more media technologies including televisions and formally proposed the Media Equation to suggest that users' responses to technologies are fundamentally social and natural [102].

✉ Kun Xu
kun.xu@ufl.edu

1 College of Journalism and Communications, University of Florida, Gainesville, FL, USA

2 Department of Media, Journalism, and Film, Miami University, Oxford, OH, USA

At the core of users' social responses to technologies is the role of social cues in constituting the characteristics of these technologies. Social cues are defined as "biologically and physically determined features salient to observers because of their potential as channels of useful information" [30, p. 2]. In interpersonal communication, social cues include but are not limited to one's eye gaze, appearance, facial expressions, movements, and gestures. These cues have been widely adopted and designed into social robots to enhance their usability and communication efficiency. For example, the widely used social robot *NAO*, developed by Aldebaran Robotics, is equipped with eye gaze, hand movements, language abilities, and synthetic voices. The zoomorphic robot *Spot*, invented by Boston Dynamics, is designed with a dog-like shape and capabilities of walking and running. Furthermore, Hiroshi Ishiguro's android *Geminoid* serves as a prototype of the replica of Ishiguro himself and consists of a constellation of subtle non-verbal cues such as frowns, blinks, and haptic feedback.

Combinations of social cues may be translated into social signals, which are referred to as the meaningful interpretations of the social cues [30]. Examples of social signals include perceived emotions, personalities, identities, competency, interactivity, and so on. For instance, the control of movement speed and frequency can constitute users' perception of the personality of a social robotic technology [66]. Therefore, to understand how social robots exhibit the social signals such as personalities and emotions, and how these social signals influence users' social responses, it is crucial to first uncover the power of social cues that construct the human features of these social robotic technologies.

While the effects of social cues have been documented in research on various technologies such as computers, chatbots, voice assistants, and telepresence robots [9, 49, 134], limited research has directly responded to Nass and Moon's call for studies on whether some social cues are more powerful than others in eliciting social perception, attitudes, and behavior [87]. As was argued by Nass and Moon [87], "perfect implementations of technologies mimicking human characteristics may generate powerful social responses, but it is not all clear whether ersatz versions of these technologies are close enough to elicit more social responses than would have occurred in their absence" (p. 97). Although there has been some research that distinguishes the different effects between voice and text [20], gestures and movements [136], and human voices and synthetic voices [41], there lacks a systematic organization of the existing knowledge about the distinct effect of each social cue in the context of human–robot interaction (HRI). Therefore, this study aggregates prior research findings and conducts meta-analyses to understand the effects of social cues on users' social responses toward social robots. Specifically, two meta-analyses were conducted to examine the strengths of the effects of social cues on users' social presence with and trust in social robots. The meta-analytic approach would be useful here as it resolves the ambiguity about the research status quo by proffering a method for combining the existing scholarly findings [107].

The results of the meta-analyses can make theoretical, practical, and methodological contributions to HRI. On the theory level, recognizing the effect size of each single social cue could contribute to a hierarchy of the power of social cues. This hierarchy could lead researchers to understand how various dimensions of social robotic technologies differ in their impacts on users' social reactions [87]. Furthermore, past works have sought to build stronger theoretical frameworks to extend the CASA paradigm, which include the attempts to distinguish social cues based on their distinct potential for activating users' social responses [74]. The results of the meta-analyses would thus validate these propositions and further refine and update the CASA paradigm.

On the practical level, technology developers will have a comprehensive understanding of the power of each single cue when designing social robots. With the results of the meta-analyses, developers may prioritize the design of certain cues to not only manage the budget of product promotion and development but also strategically maximize the technology credibility and acceptance.

Methodologically, HRI researchers often need to manipulate the social cues of social robots to achieve desired differences between treatment and control conditions. The effect sizes of social cues could inform researchers' decision-making in creating variances in experimental conditions. When planning new studies, researchers can also refer to these effect sizes to calculate desired sample sizes [59].

2 Literature Review

2.1 The CASA Paradigm and Human–Robot Interaction

The early CASA paradigm proposes that humans respond to computers as if they were social actors. Based on a series of studies on computers and televisions, Reeves and Nass [103] proposed the Media Equation to indicate that users' interactions with a range of technologies are fundamentally social and natural. As an example of the CASA study, Nass and colleagues [85] labeled a television that only played either news or entertainment programs as a specialist and another television that played both news and entertainment programs as a generalist. They found that participants liked the specialist television more than the generalist one and rated the specialist television as more informative and superior in their area of expertise. A more recent study indicated that users evaluated a phone that they had prior interaction with as more friendly and competent, which confirms that users transfer etiquette norms not only to computers but also to mobile phones [15]. Reeves and Nass [103] explained that individuals develop social responses to technologies because our brain has not evolved to distinguish technology-mediated experience and non-mediated experience. Based on evolutionary psychology, we interact with simulations of social actors as if they were social and real.

In the past few years, the CASA paradigm has been applied to studies on social robots. The term "social robot" has been explicated from various perspectives [12, 27]. For example, Duffy [27] highlighted the physical embodiment nature of a social robot, defining it as "a physical entity embodied in a complex dynamic, and social environment sufficiently empowered to behave in a manner conducive to its own goals and those of its community" (p. 177). Similarly, Li et al. [71] emphasized the mechanical components of a social robot and conceptualized it as "devices with mechanical moving parts that interact in socially appropriate ways" (p. 3). However, consensus has not been achieved regarding whether a social robot should be physically present. For instance, Zhao [140] defined humanoid social robots as "human-made autonomous entities that interact with humans in a human-like way" (p. 405). Similarly, Fox and Gambino [33] referred to social robots as "human-made technologies that can take physical or digital form, resemble people in form or behavior to some degree, and are designed to communicate with people" (p. 295). Both definitions specify that social robots do not necessarily need to be physically present; rather they can assume the forms of virtual agents, conversational agents, chatbots, or voice assistants.

Other scholars put more emphases on social robots' capacities for social interactions. For instance, Shin and Choo [115] argued that social robots should interact with humans according to some social rules and social behavior. Bartneck et al. [8] defined a social robot as "an autonomous or semi-autonomous robot that interacts and communicates with humans by following the behavioral norms expected by the people with whom the robot is intended to interact" (p. 592). Lee et al. [63] suggested that the primary characteristic of a social robot is to afford interactions with humans. While scholars' definitions of social robots may diverge on whether they should be physically present, there have been common characteristics among these definitions. That is, social robots need to at least feature a certain degree of automation and possess the capacity for social interactions with humans.

As some examples of the application of the CASA paradigm in social robot research, Horstmann et al. [53] looked at participants' reactions when they were given the choice to switch off the robot *NAO*. They found that participants were more likely to let the robot stay switched on when the robot verbally objected to being switched off. This study indicated that participants perceived a robot that demonstrates both autonomy and opposition as more human-like and accepting. Stock-Homburg et al. [117] examined the effect of conversational flow in HRI at workplace. Their study indicated that the android robot *Elenoide* was perceived to demonstrate greater conversational flow than the humanoid robot *Pepper*, given that the android robot *Elenoide* had a higher degree of human-likeness in generating socially engaged and interactive scripts than *Pepper*.

2.2 Social Cues

As the CASA paradigm was proposed over two decades ago, in the past few years scholars have started to expand, refine, and update the CASA paradigm to make it more theoretically tenable and testable. One approach to expand the CASA paradigm is to center on the role of social cues. Whereas much CASA literature has cast light on the effects of single social cues, or the accumulative effects of social cues [e.g., 5, 13, 37, 83, 93, 121], limited research has systematically clarified how individual social cues may exert unique influence on users' social responses to robotic technologies. In other words, it remains unknown whether there exists a hierarchy of social cues that reflects their distinct power over users' social responses. In prior literature, Lombard and Xu [74] argued that scholars should distinguish primary social cues from secondary ones based on users' evolution-based responses to media technologies. Specifically, primary cues (e.g., human voice, human shape) are sufficient but not necessary in evoking users' social responses, whereas secondary cues (e.g., text, machine-sounding speech) are neither sufficient nor necessary in evoking users' social responses. Compared to secondary cues, primary cues are more natural, powerful, intuitive, and salient to users' perception of socialness. The idea of concentrating on the quality of social cues was also mentioned in Reeves and Nass' [102] research, where they noted that sight and sound dominate human perception compared to other senses such as smell and touch. The cognitive miser theory also corroborated that the degree to which an individual perceives a social actor is contingent upon the quality of visual cues [31]. It was found that certain facial features such as hairstyles are sufficient to activate users' social perception [76]. All these studies imply that some cues may be more powerful than others in evoking users' natural and instinctive responses to social actors. However, past research has primarily focused on whether designing more social cues would evoke stronger social responses. Limited research has concentrated on the quality of single social cues. Therefore, to systematically parse out the individual effects of social cues, a meta-analysis of the prior findings regarding the magnitudes of the effects of social cues can contribute to researchers' understanding of the hierarchy of social cues. Below, based on prior literature, we list some major social cues that have been studied in the contexts of HRI. Although it is impossible to exhaust the literature on the effects of social cues, we seek to demonstrate the major social cues that have been designed into social robots.

2.2.1 Voice

Manipulating vocal cues has been one of the major strategies used in HRI experiments. A human voice is often preferred to a machine voice as it is perceived to be a natural and powerful modality [84]. For example, Cherif and Lemoine [18] found that users judged a virtual robotic agent using a human voice to be more credible than one using a synthetic voice. Xu [136] also found that the human voice of the social robot *Alpha* was more likely to evoke users' trust than the machine voice of the robot. Manipulating vocal cues can further help researchers contrive perceived personalities or identities. As an example, Torre et al. [123] found that participants perceived a robot with a standard British vocal accent to be more credible than those with other regional accents, suggesting that participants developed their first impression of the robot based on its perceived identity.

The advance of text-to-speech technologies has now allowed synthetic voices to feature paralinguistic cues such as tones, accents, and vocal fillers [39, 139], which has led to research that challenges the findings about the inferiority of machine voices to human voices. For instance, Gong and Lai [41] found that compared to mixed speech that combined both human and synthetic voices, a merely synthetic voice was considered as more pleasant and articulate. Another study suggested that when a computer agent's voice and appearance did not match (e.g., a machinelike face paired with a human voice, or a humanlike face paired with a machine voice), participants were less likely to disclose private information to the agent. Only when the agent presented a congruent voice and appearance (e.g., a machinelike face paired with a machine voice) did participants report more trust in the agent [42]. Therefore, even if some research has shown that human voices are more favorable, in some contexts, synthetic voices may trigger stronger effects on users' social responses. Therefore, the current meta-analyses may help discover the true effects of the vocal cues in past HRI research.

2.2.2 Appearance

Comparing humanlike and machinelike appearances is another theme in the manipulation of robots' social cues [7]. Abubshait and Wiese [2] found that compared to robot-like agents, humanlike agents were more likely to be perceived by users as having minds. Hinds et al. [48] also found that participants felt less responsible for a task when collaborating with a human-like robot than a machine-like robot, indicating that participants attributed more trust and closeness to the humanlike one than the machine-like one.

Paralleling the inconsistent findings about the vocal cues in HRI literature, humanlike appearances do not always lead to stronger social responses than machinelike appearances. The "uncanny valley" experience is a sign of users switching from high acceptance to revulsion when a technology appears highly humanlike but fails to perfectly mimic human appearances [81], notwithstanding that its postulation is still being tested [44]. To take another example, Lee [62] found that although virtual agents with anthropomorphic cartoon images were perceived as more attractive compared to text-based agents, they did not produce stronger flattery effects. Therefore, even though research suggests that humanlike appearance is closely associated with mind attribution [6, 77], these mixed findings entail the aggregation of existing findings about the effects of appearances in HRI.

2.2.3 Movements

Past research has suggested that humans have evolution-based responses to the motions of objects. The Heider-Simmel [46] experiment suggests that humans spontaneously attribute mental characteristics to animated geometric shapes such as triangles and circles. Johansson's [54] experiment on users' perception of bright dots also indicates that humans can automatically identify human gestures even through a limited number of spots that represent body joints. Humans' sensitivity to movements has been further found in HRI studies where, for example, Li and Chignell [69] tested the effects of robot gestures and found that a bear-like robot's simple arm and head movements (i.e., without moveable fingers, wrists, or elbows) can evoke participants' emotional responses and the perceived likeability of the robot. Similarly, Salem et al. [109] found that a robot's gestural behavior led to more social perception and future contact intentions, especially when the meanings of the gestures were aligned with the robot's verbal messages.

While the aforementioned research attests to the strong effects of robot movements on users' social responses, some research has called for a closer investigation into movement cues such as motion frequency, motion speed, or motion patterns. For instance, Morewedge et al. [80] noted that people tend to perceive robots as social actors when the robots move at a similar speed to human movement. Xu [136] found that gestural movements had stronger effects on perceived attraction and future use intentions than random movements. Considering that movements have received growing attention in HRI [63, 121, 124], it is time that scholars gathered prior findings to understand the true effects of kinetic cues in HRI.

2.2.4 Language

To increase communication efficiency and effectiveness, researchers have further added language styles including powerful versus powerless language [3], dominant versus compliant language [86], and anthropomorphic versus non-anthropomorphic language [108] to robotic technologies. Choi et al. [21] compared literal language and figurative language on a service robot and found that people responded favorably toward a service robot using literal language. Literal language further increased the credibility of the robot during its interaction with humans. Hoffman et al. [51] manipulated language to be perceived as warm or cold and found that the robot *Pepper* using warm language was rated as more likable and that the perceived warmth counterbalanced the negative influence of the robot's erroneous behavior.

In addition to voice, appearance, and movement, and linguistic cues, scholars have used other cues to manipulate social robots' perceived mental, emotional, and physical status. These cues include but are not limited to facial expressions [40, 125], eye gaze [4, 97], haptic cues [70], distance [67], and olfactory cues [17]. As an example, Kim et al. [57] replicated the Hawthorne effect in HRI and found that a robot designed to demonstrate eye gaze was more effective in leading participants to feel moral pressure and motivating voluntary donation than a robot without such design. Thus, in our meta-analyses, we include these social cues and expect to draw conclusions based on existing research.

2.3 Social Responses

Users' social responses to technologies include their social perception, social attitudes, and social behavior [87, 136]. This study focuses on two dimensions of users' social responses to social robots: social presence and trust. Specifically, social presence serves as an indicator of users' social perception, while trust reflects one type of users' social attitudes. Meanwhile, in prior research, social presence has been considered as first-degree social response and trust as second-degree social response [66], as social presence is more related to the identification and interpretation of the fundamental social dimensions presented by a technology, and trust is more related to the attitudinal response that occurs after the recognition of the social dimensions.

These two concepts are related but different, given that second-degree responses may not always arise concurrently with first-degree responses [66]. This study conducts two separate meta-analyses on the relationships between social cues and social responses. One meta-analysis centers on the effects of social cues on users' social presence experience and the other centers on the effects of social cues on users' trust in social robots. Given the space limit, we did not include other types of social responses in our meta-analyses such as users' future use intention, perceived attraction, or perceived competency. That said, researchers may continue to aggregate findings about the effects of social cues on various types of social responses.

2.3.1 Social Presence

Presence research has spanned virtual reality, augmented reality, HCI, and HRI. Presence refers to the experience in which individuals fail to perceive or acknowledge the role of a technology in their communication environment and respond as if the technology was not there [73]. In other words, presence occurs when individuals perceive the technology-mediated experience as non-mediated. Past research has categorized presence into two types: spatial presence and social presence. Whereas spatial presence generally denotes the sense of "being there" [45], social presence affords the sense of "being with another" [10, p. 456]. To be more precise, Lee [61] conceptualized social presence as "a psychological state in which virtual (para-authentic or artificial) social actors are experienced as actual social actors in either sensory or non-sensory ways" (p. 44).

When explicating social presence, scholars have noticed that users may not only perceive media characters within technologies as social actors, but also perceive the technologies per se as social actors [73]. For example, interactions with computers, humanoid social robots, automated teller machines, and smart speakers all involve such experience of (at least partially) perceiving technologies as social entities. In these interaction processes, users may feel as if they were interacting with intelligent social beings, form the perception of being with another person, and respond to them as if they were real [10, 61, 73]

Social presence has been viewed as an important concept in prior HRI research. For instance, Xu [136] examined the effects of the social robot *Alpha*'s gestural movements and non-gestural movements and found that those who had positive attitudes toward robots felt stronger social presence when Alpha demonstrated gestural movements, while those who had negative attitudes toward robots reported stronger social presence when presented with non-gestural movements. Furthermore, Lee, Jung, Kim, and Kim [63] found that the physical embodiment of a social robot had positive effects on social presence, which further increased users' positive evaluation of the robot. Similarly, Lee, Peng, Jin, and Yan [66] found that users' social presence of the zoomorphic robot *AIBO* can lead to its perceived social attraction and users' enjoyment of the interaction. Based on these studies, examining the effects of social cues on social presence would be meaningful as social presence has often served as a mediator between the social cues of social robots and users' subsequent cognitive, affective, or behavioral responses [63, 66].

2.3.2 Trust

Trust refers to the perceived competence or reliability of a person/agent [118]. In HRI studies, social cues have been found to exert positive effects on users' trust in social robots. Nomura and Kanda [93] found that people were more likely to perceive a robot as trustworthy when it was designed with eye gaze. In addition, receiving immediate, encouraging, and empathic feedback facilitated human–robot cooperation and augmented the perceived trustworthiness of the robot [129].

It is necessary to understand the overall effects of social cues on trust also because trust in robotic technologies can further lead to attitudinal and behavioral change. Gaudiello et al. [35] found that strengthening users' trust in a humanoid robot led to their acceptance of the robot as a decision-making assistant. You and Robert [138] found that letting robots exhibit the same workstyles as humans enhanced users' trust in these robots, which increased their willingness to work with the robot partner. Therefore, like social presence, trust plays a key role in the relationship between social cues and users' acceptance of and future use intention toward social robots. Increasing the perceived trustworthiness of social robots may lead to users' positive social attitudes toward them.

Thus far, we have reviewed HRI literature on the effects of social cues on users' social presence and trust. We expect the meta-analyses to be useful and illuminating insofar as the meta-analyses can (1) inform us about the extant knowledge in the HRI field, (2) convert statistical significance into interpretable effect magnitudes, (3) minimize wasted data and include non-significant results as part of the full picture of the HRI research program, and (4) find moderator variables that affect the effects of social cues on users' social responses to social robots. As Rosenthal and DiMatteo [107] argued, a meta-analysis can generate conclusions that are more precise and credible than those presented in a single study or a non-quantitative narrative review. While past meta-analyses or systematic reviews on HRI have revealed that robots' humanlike behavior does not always affect children's trust in social robots and it is robot performance that acts as the most important contributor to users' trust [43, 118], these studies have either only focused on the children-robot relationship or used robot attributes as a general dimension in meta-analyses. According to Naneva et al. [82], the influence of the design of social cues on people's attitudes toward robots has not been quantified or reviewed comprehensively. Thus, given the strengths of the meta-analytical approach, this study is the first to examine the aggregated effects of social cues on the first-degree social response and the second-degree social response. Specifically, we conduct two meta-analyses to examine the overall magnitude of the effects of social cues on social presence and trust in HRI contexts. We propose the following research questions in the first meta-analysis focusing on the effects of social cues on social presence.

RQ1: What are the overall effects of social cues on users' social presence?
RQ2: What are the effects of individual social cues on users' social presence?
RQ3: Which subgroup of social cues has the greatest effect on users' social presence?
RQ4: What factors moderate the effects of social cues on users' social presence?

We then propose the following research questions in the second meta-analysis focusing on the effects of social cues on trust.

RQ5: What are the overall effects of social cues on users' trust?
RQ6: What are the effects of individual social cues on users' trust?
RQ7: Which subgroup of social cues has the greatest effect on users' trust?
RQ8: What factors moderate the effects of social cues on users' trust?

3 Method

The procedure below follows the steps in Rosenthal and DiMatteo's [107] seminal work on meta-analysis and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol, which offers a checklist to help researchers conduct and report results from the meta-analyses [79].

3.1 Selection of the Literature

To identify academic articles that focus on the effects of social cues on trust and social presence, four online databases were used to search the literature: PsycINFO, EBSCO Host Communication and Mass Media Complete, Web of Science, and ACM Digital Library. The search included all the articles available in these databases up to May 13, 2020 (i.e., the ending date of this project). These databases were selected because they are the most widely used databases in psychology, communication, computer science, information science, and engineering. Similar search strategies have been used in prior research [26, 32, 96, 110]. As researchers use various terms to refer to social robots, social cues, social presence, and trust, we expanded our search scope to include all the relevant articles. The search for literature on social robots included articles about chatbots, virtual assistants, conversational agents, and other relevant technologies. The search for literature on social cues included articles about robot appearances, faces, kinetic cues, eye gaze, and so on. To include all literature on social presence, we included articles not only about social presence but also about social responses. The search for literature on trust was focused on terms related to trust and credibility. Overall, our search strategies were combined using the following Boolean expression: ("social robot*" OR "social bot*" OR machin* OR chatbot* OR "virtual assistant*" OR "computer agent*" OR "voice assistant*" OR "voice agent*" OR "conversational agent*") AND ("social cue*" OR shape OR appearance* OR fac* OR "eye gaz*" OR "eye contact*" OR gestur* OR kine* OR sound* OR voic* OR vocal OR language* OR speech* OR move* OR motion* OR "physical distance" OR spac* OR text*)

AND ("social presence" OR "social response" OR trust* OR credib*).

3.2 Inclusion Criteria

After determining the search terms, inclusion criteria were established to select articles from the search results. Each study should meet the following criteria: (1) It is a peer-reviewed journal article or conference proceeding (e.g., *CHI* proceedings, *IEEE* proceedings). (2) It includes quantitative measurement of social presence or trust. (3) It contains the manipulation of at least one social cue as the independent variable. (4) The study is conducted in the context of human-social robot interaction. Research on deep learning systems, algorithms, and the manipulation or operation of industrial robots is excluded from the analyses. (5) It includes sufficient details to calculate effect sizes. (6) It is published in English.

After applying the initial search strategies, 4368 articles were retrieved and uploaded to Covidence, an online management tool for title/abstract screening, full-text screening, data abstraction, and quality assessment. We first removed 1251 duplicated articles. Next, two trained research investigators separately checked the titles and abstracts of all the remaining 3117 articles. If their votes yielded inconsistent results, a third investigator on Covidence would vote for the qualification of the article. After this procedure, 2888 articles were disqualified, leaving 229 articles for full-text screening. The same voting process was then repeated using Covidence. Among these articles, 66 were excluded due to a lack of social cues as independent variables, along with 49 articles excluded due to a lack of social presence or trust being used as dependent variables. Additionally, a total of 46 articles were removed because they did not involve quantitative measurement of the variables, and 13 articles were removed because they were not examining social robots (e.g., studies on deep learning). Seven articles were excluded because they were not full papers. In addition, five articles were removed because their dependent variables used behavioral indicators such as resource allocation [e.g., 123], three articles were excluded because they did not provide sufficient information to determine the statistical relationships among variables, and one article was removed because the authors reported the same data in another publication. The remaining 39 articles were included for coding. The diagram for the literature search and screening is shown in Fig. 1.

3.3 Variable Coding

A coding scheme was developed to extract the details of the eligible studies. The coded variables included the main findings of each study along with its method (e.g., survey, experiment), within/between-subjects design, conceptualizations and operationalization of independent variables and dependent variables, valence of the relationship, sample size, group size, study settings (e.g., university, MTurk), sample demographics, mean values, standard deviation, t values, F values, correlation r, regression results, type of social robots, study tasks/contexts, sample location, and major theoretical frameworks. During this coding process, if essential data were not reported, we contacted the authors to request more information. If the authors did not respond, we excluded the studies from our sample ($n = 3$).

As various types of social cues were examined in these studies, we categorized them into six subgroups: the effects of appearance cues (e.g., humanlike robots vs. machinelike robots), facial cues (e.g., robots with eye gaze vs. robots without eye gaze), vocal cues (e.g., human voices vs. machine voices), language cues (e.g., anthropomorphic language vs. non-anthropomorphic language), movement cues (e.g., smooth movements vs. non-smooth movements), and others (e.g., touch allowed vs. no touch allowed). Details from each study were presented in Appendix 1a and Appendix 1b.

3.4 Effect Size Calculation

We referred to Rosenthal and DiMatteo's [107] recommendation of using the Pearson correlation r to compute effect sizes. While correlation r and Cohen's d can be easily converted to each other, using r has several advantages: (1) based on one degree of freedom, r allows for a representation of the continuous relationship between the independent variables and the dependent variables; (2) r is simpler to interpret when researchers need to refer to the variance of the factor effects [107].

As some studies tested both social presence and trust, we treated each dependent variable as a unit of analysis [32]. If a study reported the effect sizes, we directly used them and converted them to r. When effect sizes were not reported in the study, we searched for t values or F values with one degree of freedom, as both can easily be converted to r scores [107]. When F or t values were not provided, we searched for the group size, mean values, standard deviation of the compared conditions, computed d, and converted to r [11]. If a dependent variable was measured with different scales (e.g., affective trust and cognitive trust), their respective r values were standardized through Fisher Z transformation, averaged, and converted back to a single r [26]. If none of these values was provided, we further searched for p values or p range, as these p statistics can be converted to one-tailed standard normal deviate Z scores, which can further be transformed to r [107]. If the study only reported that the results were non-significant, we first assigned zero to the effect size of the study (i.e., zero-coded), which led to an underestimated overall effect size. We then removed this single effect size from the sample (i.e., max-coded), which led to an inflated effect size. Hence, a range of the effect

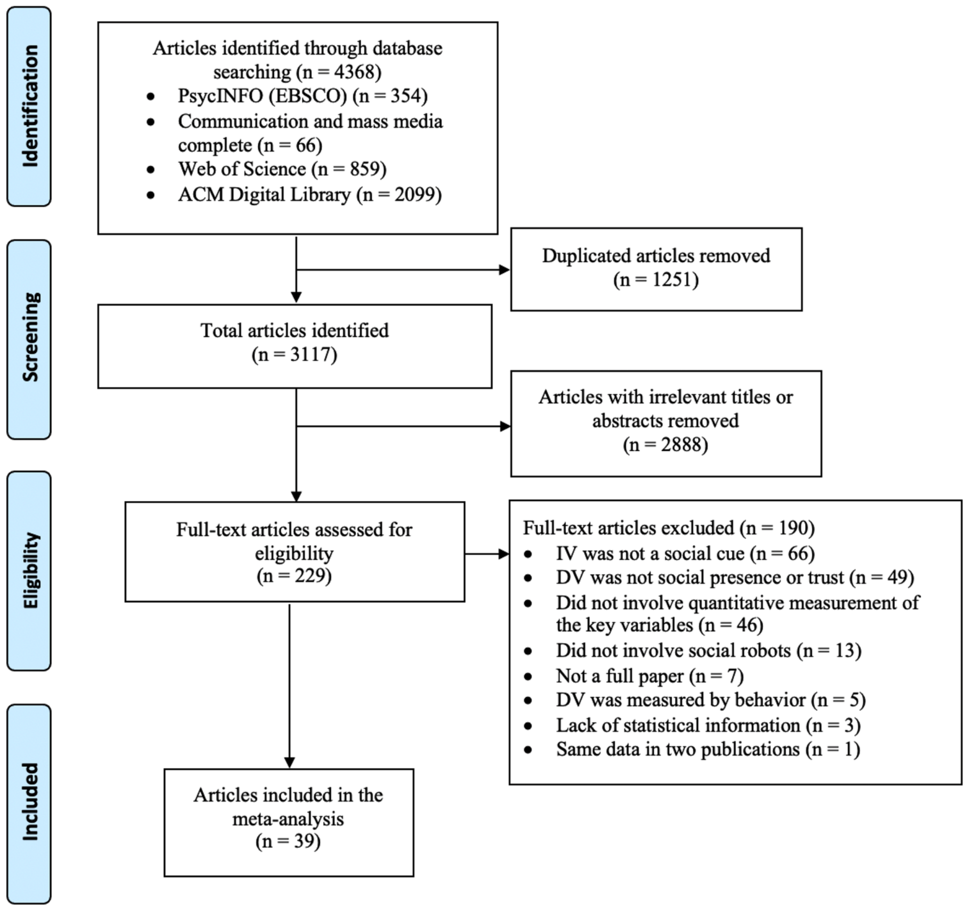


Fig. 1 PRISMA flow diagram of the literature search and screening

sizes could be obtained to account for the studies that merely reported non-significant results [106, 114].

The direction of the effect sizes depends on the comparisons of the groups. We decided that more natural, lifelike, or humanlike cues were compared against unnatural, artificial, or machinelike cues. For example, r was coded positive when humanlike social robots evoked stronger social presence than machinelike robots, whereas it was coded negative when machinelike robots exerted stronger influence. Following the Johnson and Eagly [55] perspective that if multiple levels of comparisons are made in a study, one could compare the means of the high and the low levels, in our meta-analyses, if multiple levels of social cues were manipulated in a study, we compared the most humanlike manipulation against the most machinelike one. Specific comparisons were provided in Appendix 1.

After correlation r was calculated for each effect size, we transformed the value into a Fisher Z score to normalize the distribution [24]. Then the unweighted value Z_r was weighted by $N - 3$ in each study to ensure that studies with larger sample sizes received greater weight. Next, both unweighted Z_r and weighted Z_r were averaged and converted back to r to form unweighted effect sizes and weighted effect sizes for all studies and each subgroup of studies. The random effects confidence interval was calculated based on the unweighted mean of Z_r [107].

3.5 Effect Size Interpretation and Analyses

We referred to Cohen's [22] convention that a correlation coefficient of 0.10 should be interpreted as a small effect size, 0.30 as a medium effect size, and 0.50 as a large effect size. We also conducted analyses of how each subgroup effect size differed from the others. Two independent subgroup effect size r values and their respective sample sizes were converted to a Fisher Z score [23]. The value and the significance of the Z score indicate which effect size is significantly larger than the other based on two different sample sizes.

We conducted heterogeneity tests based on random-effects models for overall effect sizes and each subgroup effect size [11]. Random effects models allow for the assumption that the studies are randomly sampled from a population of studies and can be generalized to this population [55]. Q-statistics were computed using the R package *metaphor*. The Q-statistic is a Chi-squared test with $k - 1$ degrees of freedom, which reflects the heterogeneity of variance in the effect sizes. A significant result on the Chi-squared test indicates that the true effect size variance was not only caused by sampling error but also by potential moderators. To quantify the percentage of the true effect size that is attributed to the between-study variance, I^2 was computed. I^2 describes the percentage of total variance across studies that is due to the inconsistency of effects rather than chance [47]. If I^2 is greater than 75%, considerable heterogeneity is present. If I^2 is below 40%, the studies in the analysis are homogeneous [24]. I^2 is 0% if there is no observed heterogeneity [47]. Meta-regression was used to calculate the effects of moderators on effect sizes.

4 Results

As some articles include multiple studies, among the 39 articles included for the meta-analyses, a total of 25 studies ($N = 2498$) examined the relationship between social cues and social presence. A total of 44 studies ($N = 4147$) examined the relationship between social cues and trust. To examine the overall effects of social cues on users' social presence and trust, both unweighted effect size r and weighted effect size r were computed. Below we report the weighted effect sizes.

4.1 Meta-Analysis 1: Effects of Social Cues on Social Presence in HRI

To examine the overall effects of social cues on social presence in HRI (*RQ1*) and the effects of individual social cues (*RQ2*), the meta-analysis suggested that the social cues of social robots had small-sized effects on users' social presence ($0.17 \leq r \leq 0.18$). The subgroup analyses revealed that the appearances of the social robots had very limited effects on social presence ($0.07 \leq r \leq 0.08$). Manipulating the voice ($r = 0.16$) and language ($r = 0.12$) had small-sized effects on social presence. Additionally, movement cues had small-to-medium-sized effects on social presence ($0.24 \leq r \leq 0.30$). Facial cues had large-sized effects ($r = 0.69$), and other types of cues ($r = 0.45$) including haptic cues had medium-sized effects on social presence. Results are shown in Table 1.

To understand which subgroup of social cues had the largest effect on social presence (*RQ3*), Fisher Z scores were calculated to examine whether the effect sizes of individual social cues were significantly different from each other. Results suggested that with regard to social presence, facial cues had the largest effect size, followed by movement, voice, language, and appearance cues. Statistically, facial cues had significantly larger effect sizes than any other category of social cues. However, considering the small sample of the study on facial cues ($k = 1$), movement cues merit more attention as they had significantly larger effects than voice, language, and appearance cues on users' social presence (see Table 2).

To understand what factors moderated the effects of social cues on social presence (*RQ4*), heterogeneity tests were conducted using the Q-statistic and I^2. A Chi-square test of the Q-statistic suggested that additional moderators affected the true effect size variance. As shown in Table 1, the heterogeneity tests for the overall effects of social cues on social presence were significant, $p < 0.001$. Among the subgroup social cues, only the effects of language cues on social presence were homogeneous.

Potential moderators were searched to examine if they account for the between-study variance of the true effect sizes. Based on prior literature [106], year of publication (i.e., recency of the publication time), the mean age of the sample, gender distribution, the embodiment of the social robots, the physical presence of the robots, sampling characteristics (i.e., whether the participants were recruited from universities), and task properties (i.e., whether participants were asked to interact with the technologies) were included as moderators. Results from meta-regression suggested that the selected moderators did not affect the relationship between social cues and social presence.

4.2 Meta-Analysis 2: Effects of Social Cues on Trust in HRI

To examine the overall effects of social cues on users' trust (*RQ5*) and the effects of individual social cues (*RQ6*), the meta-analysis suggested that overall, the social cues had small-sized effects on users' trust in social robots ($0.14 \leq r \leq 0.17$). The subgroup analyses suggested that manipulating social robots' appearances ($0.10 \leq r \leq 0.12$), voices ($r = 0.13$), language ($r = 0.22$), and facial cues ($0.17 \leq r \leq 0.28$) had small-sized effects on social robots. By contrast, movement cues ($0.20 \leq r \leq 0.31$) and other types of cues ($0.08 \leq r \leq 0.34$) had small-to-medium-sized effects on users' trust in social robots. Results are shown in Table 3.

To understand which subgroup of social cues had the largest effect on trust (*RQ7*), Fisher Z scores were calculated to examine whether the effect sizes of the individual social cues were significantly different from each other. Results suggested that using the lower boundary of the effect sizes (i.e., when effect sizes were zero-coded), language had the largest effect on users' trust in social robots, followed by movements, facial cues, voices, and appearances. However, the effect of language cues was not significantly larger than the movement cues. Using the upper boundary of the effect sizes (i.e., when effect sizes were max-coded), the analyses suggested that movement cues had the largest effect, followed by the effects of facial cues, language, voice, and appearance cues. The effect of movement cues was also not significantly larger than the facial cues. Combining zero-coded and max-coded

Table 1 The meta-analysis results for the effects of social cues on social presence

Independent variables	K	Unweighted r	Weighted r	95% CI	N	Q	I^2 (%)
Social Cues (zero coded)	25	.272	.173	[.162, .375]	2498	104.81***	85.75
Social Cues (max coded)	23	.294	.180	[.179, .401]	2402	101.95***	86.85
Appearance (zero coded)	4	.106	.065	[− .341, .513]	239	25.79***	91.22
Appearance (max coded)	3	.140	.080	[− .470, .660]	191	25.56***	94.20
Voice	7	.195	.160	[.079, .306]	1215	15.72*	68.30
Language	5	.137	.119	[.054, .218]	627	4.05	.02
Movement (zero-coded)	4	.331	.236	[− .066, .638]	230	25.70***	87.90
Movement (max-coded)	3	.429	.295	[− .034, .742]	182	22.41***	88.46
Facial cues	1	.693	.693	[.402, .857]	24	N/A	N/A
Others	4	.508	.448	[.343, .642]	163	3.78	17.57

K: the number of effect sizes; CI*:* confidence interval; *Q*-statistics: Chi-squared scores with $K - 1$ degrees of freedom; I^2: the percentage of variation across studies that is due to heterogeneity. Rows that do not differentiate zero-coded versus max-coded mean that no studies in the subgroup need zero or max coding

$^*p < .05$; $^{***}p < .001$

Table 2 Comparisons of the effect sizes of social cues on social presence

	Appearance (0)	Appearance (max)	Voice	Language	Movement (0)	Movement (max)	Facial cues
Appearance (0)	N/A						
Appearance (max)	N/A	N/A					
Voice	− 1.35	− 1.04	N/A				
Language	− .71	− .47	.85	N/A			
Movement (0)	− 1.89*	− 1.63	− 1.09	− 1.56	N/A		
Movement (max)	− 2.41**	− 2.14*	− 1.78*	− 2.18*	N/A	N/A	
Facial cues	− 3.46***	− 3.36***	− 3.15**	− 3.31***	− 2.69**	− 2.38**	N/A

(0): the cue was zero coded. (max): the cue was max coded. Each value represents a Z score computed using the weighted *r* and sample size of each feature. Whether zero-coded or max-coded, the effects of cues are ranked as follows: facial cues > movement cues > voice > language > appearance

$^*p < .05$; $^{**}p < .01$; $^{***}p < .001$

results, overall, language cues, movement cues, and facial cues tended to exert larger effects than voice and appearance cues (See Table 4).

To understand what factors moderated the effects of social cues on trust in HRI (*RQ8*), as shown in Table 3, the heterogeneity tests for the overall effects of social cues on trust were significant, $p < 0.001$. Only the effects of language cues and appearance cues were homogeneous. The same moderators selected for meta-analysis 1 were included for meta-regression. Results suggested that physical presence positively predicted the zero-coded effects of social cues on trust, $B = 0.26$, $p = 0.013$ and the max-coded effects of social cues on trust, $B = 0.27$, $p = 0.022$. The gender distribution negatively predicted the zero-coded effects on trust, $B = -0.82$, $p = 0.002$, and the max-coded effects on trust, $B = -1.40$, $p < 0.001$, meaning that samples with more female participants were more likely to influence the effects of social cues on users' trust. In addition, university samples negatively predicted the zero-coded effects on trust, $B = -0.30$, $p = 0.002$ and max-coded effects on trust, $B = -0.31$, $p = 0.003$, meaning that college students were less sensitive to the social cues presented by technologies. Results of the moderator analyses are shown in Table 5. Results of publication bias and forest plots are attached in Appendix 2 and 3.

5 Discussion

This study applied two meta-analyses to parse out the status quo of HRI research on the effects of social cues on two types of social responses: social presence and trust. Overall, social cues were found to have small-sized effects on both users' social presence with and trust in social robots. Although the overall effects of social cues were small, the

Table 3 The meta-analysis results for the effects of social cues on trust

Independent variables	K	Unweighted r	Weighted r	95% CI	N	Q	I^2 (%)
Social Cues (zero coded)	44	.135	.138	[.084, .185]	4147	73.15**	42.08
Social Cues (max coded)	35	.178	.173	[.120, .234]	3432	65.46***	46.66
Appearance (zero coded)	13	.055	.104	[− .041, .150]	912	15.58	1.64
Appearance (max coded)	11	.065	.12	[− .042, .172]	789	14.10	.01
Voice	8	.162	.127	[.059, .262]	1456	20.53**	68.94
Language	7	.228	.222	[.119, .331]	510	4.62	.00
Movement (zero-coded)	8	.135	.203	[− .004, .270]	567	16.89*	55.99
Movement (max-coded)	4	.266	.310	[.123, .397]	360	4.03	33.69
Facial cues (zero-coded)	6	.198	.174	[.039, .346]	473	12.75*	59.44
Facial cues (max-coded)	4	.292	.277	[.153, .419]	293	4.07	29.98
Others (zero-coded)	2	.173	.079	[− .166, .475]	99	1.98	49.60
Others (max-coded)	1	.336	.336	[− .078, .651]	24	N/A	N/A

K: The number of effect sizes; CI: confidence interval; Q-statistics: Chi-squared scores with $K-1$ degrees of freedom; I^2: The percentage of variation across studies that is due to heterogeneity. Rows that do not differentiate zero-coded versus max-coded mean that no studies in the subgroup need zero or max coding

*$p < .05$; **$p < .01$; ***$p < .001$

Table 4 Comparisons of the effect sizes of social cues on trust

	Appearance (0)	Appearance (max)	Voice	Language	Movement (0)	Movement (max)	Facial cues (0)	Facial cues (max)
Appearance (0)	N/A							
Appearance (max)	N/A	N/A						
Voice	− .55	− .16	N/A					
Language	− 2.19*	− 1.85*	− 1.90*	N/A				
Movement (0)	− 1.89*	− 1.55	− 1.58	.325	N/A			
Movement (M)	− 3.46***	− 3.13**	− 3.27**	− 1.37	N/A	N/A		
Facial cues (0)	− 1.26	− .95	− .91	.78	.48	2.06*	N/A	
Facial cues (max)	− 2.67**	− 2.39**	− 2.44**	− .80	− 1.09	.46	N/A	N/A

(0): the cue was zero coded. (max): the cue was max coded. Each value represents a Z score computed using the weighted r and sample size of each feature. When zero-coded, the effects of cues are ranked as follows: language cues > movement cues > facial cues > voice > appearance. When max coded, the effects are ranked as follows: movement cues > facial cues > language > voice > appearance

*$p < .05$; **$p < .01$; ***$p < .001$

positive valence of the effect sizes revealed that users experienced greater levels of social presence and developed more trust in social robots when the HRI experience was designed to be more humanlike, intuitive, natural, and spontaneous. Results also suggested that appearance cues, voice cues, language cues, movement cues, facial cues, and other types of cues had distinct impacts on users' social responses. Specifically, the Z scores based on the comparisons of the subgroup effect sizes revealed that facial cues had the largest effect on users' social presence, followed by movement, voice, language, and appearance cues. When it comes to users' trust in social robots, language cues, movement cues, and facial cues tended to exert larger effects than other cues.

The facial cues coded in the dataset included researchers' manipulation of both robots' facial expressions (e.g., smiles) and gaze features. While prior research has suggested that eye gaze and facial expressions can trigger users' evolution-based responses [4, 58, 97, 111, 112], considering the small number

Table 5 Moderator analyses of the effect sizes

Moderators	Effect on social presence (zero-coded) *B*	Effect on social presence (max-coded) *B*	Effect on trust (zero-coded) *B*	Effect on trust (max-coded) *B*
Age	− .02	− .02	.01	.01
Gender distribution	.33	.79	− .82**	− 1.40***
Physical presence	− .34	− .53	.26*	.27*
Embodiment	.26	.37	− .13	− .09
Interactivity	.25	.34	− .10	− .10
Sample	.07	.20	− .27**	− .31**
Recency	− .02	− .04	.01	.02
R^2	.29	.60	.48*	.60*

of studies on the effects of facial cues on social presence, more research should be done to validate the effects of facial cues in the future.

Movement cues have been manipulated in prior studies to examine the effects of gestures or smooth motion paths [9, 121]. Movement cues were found to have small-to-medium-sized effects on both users' social presence and trust. These results corroborated the Heider and Simmel [46] and the Johansson [54] experimental findings that human beings can intuitively and effortlessly attribute intentions to moving objects. As movements have been manipulated to reflect social signals such as robots' personalities or interactivity [e.g., 25, 130, 135], researchers should consider prioritizing the movement cues in their future experiment and product design.

One surprising finding from both meta-analyses suggests that appearance cues (e.g., humanlike shape, body size) imposed small-sized effects on users' social presence and trust. While much literature has suggested that anthropomorphic robots are preferred to zoomorphic or machinelike robots since they are more likely to evoke mind attribution [2, 72], the small effects of appearance cues could be attributed to four reasons. One is that participants may have experienced the uncanny valley effect in some HRI contexts. That is, participants' attachment to anthropomorphic social robots may have plunged when the robots seemed highly humanlike but failed to replicate human appearances [81]. Another reason is that some studies were testing the effects of anonymity. The Social Identity Model of Deindividuation Effects suggests that an anonymous online group environment may actually enhance participants' trust in online group members and foster their conformity to the group norms [116]. One study included in the current meta-analysis corroborated this theory. In the study, a mere text box evoked much stronger social presence than a computer agent visually represented by a stick figure [64]. The third explanation lies in the possibility of participants experiencing the machine heuristic [120]. That is, when a technology interface appears machinelike, participants may interpret the machine as objective, trustworthy, and reliable, which may counterbalance the positive effects of humanlike appearances on trust. The fourth explanation could be attributed to the dissolution of the novelty effects in HRI [34]. As people may have already been used to encountering anthropomorphic computer agents or humanlike social robots in their daily lives or through popular media portrayals, the momentum of the humanlike appearances may have dissipated.

Similar to appearance cues, vocal cues were found to exert small-sized effects on users' social responses. Although much literature has endorsed that human voice is a more favorable modality than synthetic voice in educational and online shopping settings [78, 131], the effects of human voice may not be as large as expected. Perhaps it is because synthetic voices have been designed to be increasingly humanlike in the past few years, which has diminished the discrepancies. Google *Duplex* is a representative example in that its AI technology carries out natural conversations to help users make reservations and book appointments [99].

It should also be noted that in addition to appearance, voice, language, movement, and facial cues, some other cues such as haptic cues were identified in our coding process. For example, having participants touch an agent's shoulder in a virtual environment had large-sized effects on social presence compared to the condition where no touch was allowed [52]. Touching the warm skin of a robot also evoked medium-sized effects on users' social presence [98]. Despite small samples, the results may inform future social robotics design by revealing that haptic cues had the potential to evoke medium-to-large sized effects. Future HRI collaboration may also allow users to have more opportunities to physically touch the social robots to increase their perception of the robots as social entities.

The moderator analyses suggest that a few factors may account for the heterogeneity of the variance in effect sizes.

First, the dimensions of embodiment and physical presence have been used to categorize different types of social robotic technologies. In prior literature, embodiment refers to the idea that intelligence needs to exist in the form of a body [141]. According to Pfeifer and Scheier [100], embodiment can be realized either as a physical robot or as a virtual agent. Comparatively, the dimension of physical presence means whether social robots are collocated in the same space as their users rather than merely presented in a virtual format. According to Li [68], many HRI studies comparing physical social robots and virtual agents have conflated embodiment and physical presence. Li [68] found that it was social robots' physical presence (i.e., whether the robot is physically present or virtually present) rather than embodiment (i.e., whether the robot has a form of body or not) that swayed peoples' social responses.

Following the conceptual differences, the moderator analyses in meta-analysis 2 confirmed that while the embodiment of a robot did not affect the overall effect sizes, the physical presence positively predicted the effects of social cues on users' trust, meaning that users may have attributed more reliability and credibility to the robots when the social robots were physically present and accessible. The results are consistent with the prior finding that it was the physical presence of a robot rather than its embodiment that induced users' social responses [68, 132]. The finding also implied that to best augment the effects of social cues on users' trust in future research, scholars could consider using more in-person settings which enable users to interact with real physical robots rather than virtually presented robots.

Second, the effect of gender distribution in our moderator analyses suggests that the effects of social cues on users' trust were more salient among female participants than male participants. That is, when social robots were designed with social cues, females were more likely to develop trust in these robots. The finding may have ethical implications. In particular, researchers should take caution in manipulating cues and fully inform males and females of the different potential perils, especially when social robots are used in contexts such as delivering fake news or misinformation.

In addition to physical presence and gender distribution, future research should seek more moderators to explain the heterogeneity of the variance in effect sizes. For example, factors such as users' prior robot use experiences, expectations for robot performances, and attitudes toward robots' social roles have been found to affect users' social presence experience in prior studies [56, 95] and thus could be added as moderators in future meta-analyses.

5.1 Theoretical, Practical, and Methodological Implications

On the theory level, results from the two meta-analyses can contribute to the construction of the CASA paradigm. The CASA paradigm suggests that users respond to a range of technologies with social cues in a social manner. However, it remains unknown whether some dimensions of the technologies are more powerful in evoking users' social responses than others [87]. The results of the meta-analyses herein contributes to the CASA paradigm by providing accumulated evidence for the distinct effects of social dimensions on users' social presence and trust in HRI contexts.

The hierarchy that represents the potential of social cues for evoking users' social responses further confirms the theoretical proposition that the quality of social cues plays a significant role in HRI. Specifically, scholars have proposed that individuals' social perception is determined by the quality of visual inputs available to the perceivers [31, 36, 76]. Similarly, Lombard and Xu [74] suggested that there exists a group of social cues that are more salient and central to users' perception of socialness than others. The meta-analytic findings have not only supported these extrapolations but also provided a direction for future research on the qualitative nature of social cues in HRI.

Results from the meta-analyses can further inform practical applications such as user interface design. With the hierarchy of social cues, developers can prioritize the design of facial expressions, eye gaze, and meaningful movements if their goal is to build a trustworthy robot that evokes users' perception of the robot as an intelligent social actor. Embedding haptic feedback would also lead to users' perception of a reliable and socially present robot.

Methodologically, understanding the power of each social cue may help researchers with experimental design or statistical power computing. For instance, researchers may foreground kinetic cues or facial cues if part of their experimental design goal is to increase users' social presence or trust in social robots. Additionally, referring to the effect size for each group of social cues, researchers can conduct a-priori power analyses to calculate the desired sample size for new studies.

Researchers should also fully understand the ethical risks of manipulating users' responses using these social cues. This study discovers that certain means of manipulating haptic cues [52] or facial cues [37] may lead to medium-to-large sized effects on users' social responses. Thus, it is possible that participants may be victims of deception or be susceptible to technologies that demonstrate a constellation of these social cues. Researchers should keep participants fully informed of the research purposes and the possible consequences of HRI.

6 Conclusions and Limitations

This study seeks to update and advance the application of the CASA paradigm in future HRI studies. Examining the effect sizes of social cues enables researchers to understand the magnitude of the effects in a standardized manner. Although a clear hierarchy of the effects of social cues needs corroboration from more divergent evidence (e.g., qualitative assessment, physiological measures), this study uses two meta-analyses and reveals that designing humanlike, natural, and lifelike cues is overall effective in evoking users' social responses such as users' social presence and trust in social robots. Meanwhile, among different types of social cues, facial cues and movement cues are more powerful than others in evoking users' social presence and trust in social robots. As one advantage of the meta-analysis, this study points the way toward essential theoretical and practical developments in future HRI research.

It is worth noting that the meta-analyses do not close the door to further theory construction and scrutiny [101]. One of the limitations here was that the number of effect sizes in some subgroups was not sufficient to reveal reliable results (e.g., effects of facial cues on social presence). Although meta-analytic approaches can be applied to as few as two studies [106], the results may be untenable when the number of studies is low. Therefore, future research could expand the scope of the search to include more studies on the relationship between social cues and social responses. Second, this study did not exhaust all the social cues that may affect HRI. For example, embodiment and physical presence were investigated as moderators in the current meta-analyses. Future research could systematically include these two cues and other cues such as haptic ones and olfactory cues to advance our understanding of HRI. Third, a future meta-analytic approach could compare the effect sizes of social signals that are composed of these social cues. For example, researchers could explore whether there exists a hierarchy that ranks the effects of the perceived personalities, identities, operation flexibilities, gender stereotypes, and other social signals. Such knowledge may further help researchers to explore the best strategies to design social robots that are efficient in communication and acceptable in our daily lives.

Funding No funding was received to assist with the preparation for this manuscript.

Availability of Data and Materials The datasets generated during the current study are available from the corresponding author on reasonable request. Some data have already been included in this article (Appendices).

Code Availability Available upon request.

Declarations

Conflict of interest There is no conflict of interest to disclose and this manuscript is not under consideration for publication at other outlets.

Ethical Approval This is a meta-analysis study. The Research Ethics Committee of the university has confirmed that no ethical approval is required.

Appendix 1a

See Table 6.

Table 6 Descriptive summary of sample studies on social presence

Study	Year	*Zr*	*N*	Name of DV	Operationalization of the technology	Sample location	Comparisons
Vocal cues							
Abdulrahman et al.	2019	0.037	118	Co-presence	Virtual human Sarah	Australia	Human versus synthetic voice
Chérif and Lemoine	2019	0.148	640	Perceived social presence	Virtual assistant: Prosper	Unknown	Human versus synthetic voice
Cho et al.	2019	0.304	84	Human-likeness	Microsoft Cortana	United States	Voice versus text
Lee and Nass	2005	0.310	72	Social presence	Voice assistant on web interface	United States	Extroverted versus introverted voice
Lee and Nass	2005	0.483	80	Social presence	Voice assistant on web interface	United States	Extroverted versus introverted voice
Xu	2019	0.064	110	Medium-as-social actor presence	Robot Alpha	United States	Human versus synthetic voice
Xu	2020	0.037	111	Medium-as-social-actor presence	Voice assistant on smartphones	United States	Human voice versus text
Language cues							
Araujo	2018	0.078	175	Social presence	Self-developed chatbot	Netherlands and United States	Anthropomorphic versus non-anthropomorphic language
Goble and Edwards	2018	0.300	67	Social presence	Robot NAO	United States	Vocal filter versus no focal filter
Straten et al.	2020	0.188	144	Social presence	Robot NAO	Netherlands	Neutral versus transparent
Velner et al.	2020	0.035	130	Social engagement	Robot NAO	Netherlands	Mixed versus no intonation
Xu	2020	0.090	111	Medium as social actor presence	Voice assistant on smartphones	United States	Anthropomorphic versus non-anthropomorphic language
Appearance cues							
Barco et al.	2020	0.752	35	Social presence	Nao/Pleo/Cozmo	Netherlands	Anthropomorphic versus zoomorphic
Lee and Nass	1999	− 0.417	48	Social presence	Web interface	United States	Stick figure versus text agent
Li et al.	2010	0.089	108	Engagement	Lego Mindstorm robot	China	Humanlike versus machinelike
Terzioğlu et al.	2020	Zero/max coded	48	Social presence	Cobot	United States	Gripper vertical, glass mounted versus gripper horizontal, no glass
Movement cues							
Terzioğlu et al.	2020	Zero/max coded	48	Social presence	Cobot	United States	Smooth versus non-smooth motion
Terzioğlu et al.	2020	0.861	48	Social presence	Cobot	United States	Idle versus no idle
Terzioğlu et al.	2020	0.482	24	Social presence	Cobot	United States	Elbow down versus elbow up

Table 6 (continued)

Study	Year	*Zr*	*N*	Name of DV	Operationalization of the technology	Sample location	Comparisons
Xu	2019	0.034	110	Medium-as-social-actor presence	Robot Alpha	United States	Gestural versus non-gestural movement
Facial cues							
Terzioğlu et al.	2020	0.853	24	Social presence	Cobot	United States	Gaze versus no gaze
Other cues							
Hoppe et al.	2020	0.813	19	Co-presence	An artificial hand	Germany	Touch versus no touch
Lee et al.	2006	0.666	32	Social presence	Robot Aibo	United States	Physical embodied versus disembodied
Lee et al.	2006	0.374	32	Social presence	Robot Aibo	United States	Physical embodied versus disembodied agent
Park and Lee	2014	0.386	80	Social presence	Pleo Robot	South Korea	Warm skin versus cool skin

Although the names of dependent variables vary, the actual measurements of these DVs were all about social presence. Zero/max coding means that as the authors only reported non-significant result without further details, the effect sizes were zero/max coded to create a range

Appendix 1b

See Table 7.

Table 7 Descriptive summary of sample studies on trust

Study	Year	*Zr*	*N*	Name of DV	Operationalization of the technology	Sample location	Comparisons
Voice cues							
Abdulrahman et al.	2019	0.065	118	Trustworthiness	virtual human Sarah	Australia	Human versus synthetic voice
Chérif and Lemoine	2019	0.068	640	Trustworthiness	virtual assistant Prosper	Unknown	Human versus synthetic voice
Chiou et al.	2020	0.334	89	Trust	online virtual human	Unknown	Human versus machine voice
Gong and Nass	2007	− 0.016	80	Trust	computer agent Baldi	United States	Human versus humanoid voice
Law et al.	2020	0.135	198	Trust	Willow Garage PR2 robot	Unknown	Voice versus text
Torre et al.	2020	0.415	110	Trustworthiness	Pre-programmed computer agent	United Kingdom	Smiling voice versus neutral voice
Xu	2019	0.283	110	Trust	robot Alpha	United States	Human voice versus synthetic voice
Xu	2020	0.024	111	Trust	Voice assistant on smartphone	United States	Human voice versus text
Language cues							
Calvo et al.	2020	0.062	42	Trustworthiness	Furhat robot	Sweden	Persuasive versus neutral language
Ghazali et al.	2019	0.300	21	Trusting beliefs	SociBot	Netherlands	Feedback versus no feedback
Hoffmann et al.	2020	0.341	81	Affective and cognitive trust	Robot Pepper	Germany	Error-free versus erroneous
Hoffmann et al.	2020	0.151	81	Affective and cognitive trust	Robot Pepper	Germany	Warm versus non-warm language
Hoegen et al.	2019	0.352	30	Trustworthiness	Portable Bluetooth Speaker	United States	Matching language style versus control
Straten et al.	2020	0.289	144	Trust	Robot NAO	Netherlands	Transparent language style versus control
Xu	2020	0.128	111	Trust	Voice assistant on smartphone	United States	Anthropomorphic versus non-anthropomorphic language
Appearance cues							
Burgoon et al.	2000	0.075	20	Credibility	Computer agent	Sweden	Picture of agent versus no picture
Castro-González et al.	2016	− 0.067	34	Trustworthiness	Baxter robot	Unknown	Full body versus one arm
De Visser et al.	2016	− 0.248	20	Trust	Automated agent	United States	Human versus agent
De Visser et al.	2016	− 0.045	17	Trust	Automated agent	United States	Human versus agent
De Visser et al.	2016	− 0.204	20	Trust	Automated agent	United States	Human versus agent
Erebak and Turgut	2019	0.115	102	Trust	Robot AILA	Turkey	Humanoid versus android robot

Table 7 (continued)

Study	Year	*Zr*	*N*	Name of DV	Operationalization of the technology	Sample location	Comparisons
Gong and Nass	2007	0.019	80	Trust	computer agent Baldi	United States	Human versus humanoid
Li et al.	2010	0.081	108	Trust	Lego Mindstorm robot	China	Humanlike versus machine like
Natarajan and Gombolay	2020	Zero/max coded	75	Trust	Nao and Sawyer robots	United States	Humanlike versus mechanical
Terzioğlu et al.	2020	Zero/max coded	48	Trust	Cobot	United States	Gripper vertical, glass mounted versus gripper horizontal, no glass
VanVugt et al.	2009	0.245	80	Perceived ethics	Computer agent	Netherlands	Correlation between body size and trust
VanVugt et al.	2009	0.161	270	Perceived ethics	Computer agent	Netherlands	Correlation between body size and trust
Weitz et al.	2019	0.587	30	Trustworthiness	Virtual agent: Gloria	Germany	Humanlike agent versus no agent
Movement cues							
Bevan and Fraser	2015	0.211	60	Trustworthiness	Robot Nao	United Kingdom	Handshake and feedback versus handshake
Castro-González et al.	2016	0.193	34	Trustworthiness	Baxter robot	Unknown	Smooth versus mechanistic movement
Terzioğlu et al.	2020	Zero/max coded	48	Trust	Cobot	United States	Smooth versus non-smooth motion
Terzioğlu et al.	2020	Zero/max coded	48	Trust	Cobot	United States	Idle versus no idle
Terzioğlu et al.	2020	Zero/max coded	24	Trust	Cobot	United States	Elbow down versus elbow up
VandenBrule et al.	2014	0.441	156	Trust	TWENDY robot	Netherlands	Motion fluency: smooth vs, shake
VandenBrule et al.	2014	0.000	87	Trust	scoreboard robot	Netherlands	Motion fluency: smooth vs, shake
Xu	2019	0.245	110	Trust	robot Alpha	United States	Gestural versus non-gestural movement
Facial cues							
Elkins and Derrick	2013	0.158	88	Trust	Embodied conversational agent	United States	Smile versus neutral facial expression
Ghazali et al.	2018	0.473	72	Trusting beliefs	SociBot	Netherlands	Upturned eyebrows and lips versus nasolabial deepener, lip corner depressor, lips toward each other
Nomura and Kanda	2015	0.240	93	Trust	Robovie-R2 robot	Japan	Normal gaze versus refrained gaze
Shamekhi et al.	2018	0.331	40	Trust	Telepresence robot	United States	Face versus no face
Terzioğlu et al.	2020	Zero/max coded	24	Trust	Cobot	United States	Gaze versus no gaze
VandenBrule et al.	2014	Zero/max coded	156	Trust	TWENDY robot	Netherlands	Gaze following movement versus static

Table 7 (continued)

Study	Year	*Zr*	*N*	Name of DV	Operationalization of the technology	Sample location	Comparisons
Other cues							
Looije et al.	2010	0.349	24	Trustworthy	iCat robot	Netherlands	Physical versus virtual
Natarajan and Gombolay	2020	Zero/max coded	75	Trust	Nao	United States	Physical versus virtual

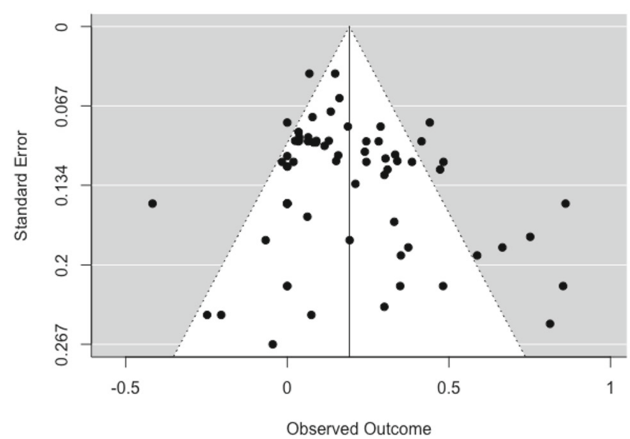
Although the names of dependent variables vary, the actual measurements of these DVs were all about trust. Zero/max coding means that as the authors only reported non-significant result without further details, the effect sizes were zero/max coded to create a range

Appendix 2

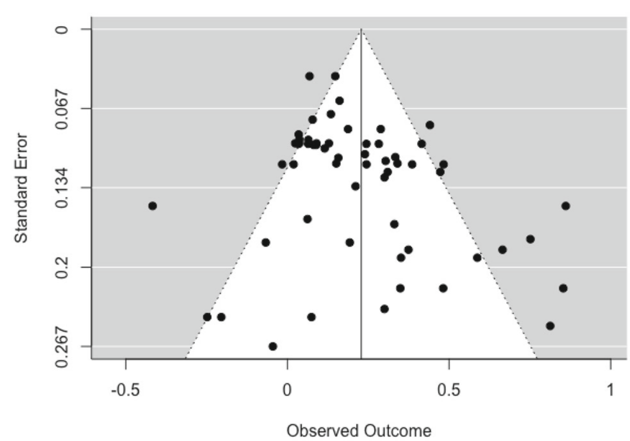
Publication Bias

To identify a potential "file-drawer-problem" that may lead to the overestimation of the overall effect of pooled relationships, a common method to detect publication bias is Rosenthal's [104] approach of fail-safe *n* calculation. This approach is to compare the value of "fail-safe *n* bias" and "fail-safe *n*." If the fail-safe *n* is larger than the fail-safe *n* bias, then the meta-analysis features no publication bias. By contrast, if the fail-safe n is smaller than the fail-safe *n* bias, it indicates a potential bias of the meta-analysis. In this study, R package *metaphor* revealed that the fail-safe *n* is 4066. Under the zero-coded condition, the fail-safe n bias was 355. Under the max-coded condition, the fail-safe *n* bias was 295. In both conditions, the fail-safe *n* bias was smaller than the fail-safe *n*, meaning that this meta-analysis presented no publication bias (see funnel plots).

Funnel plot for zero-coded meta-analysis



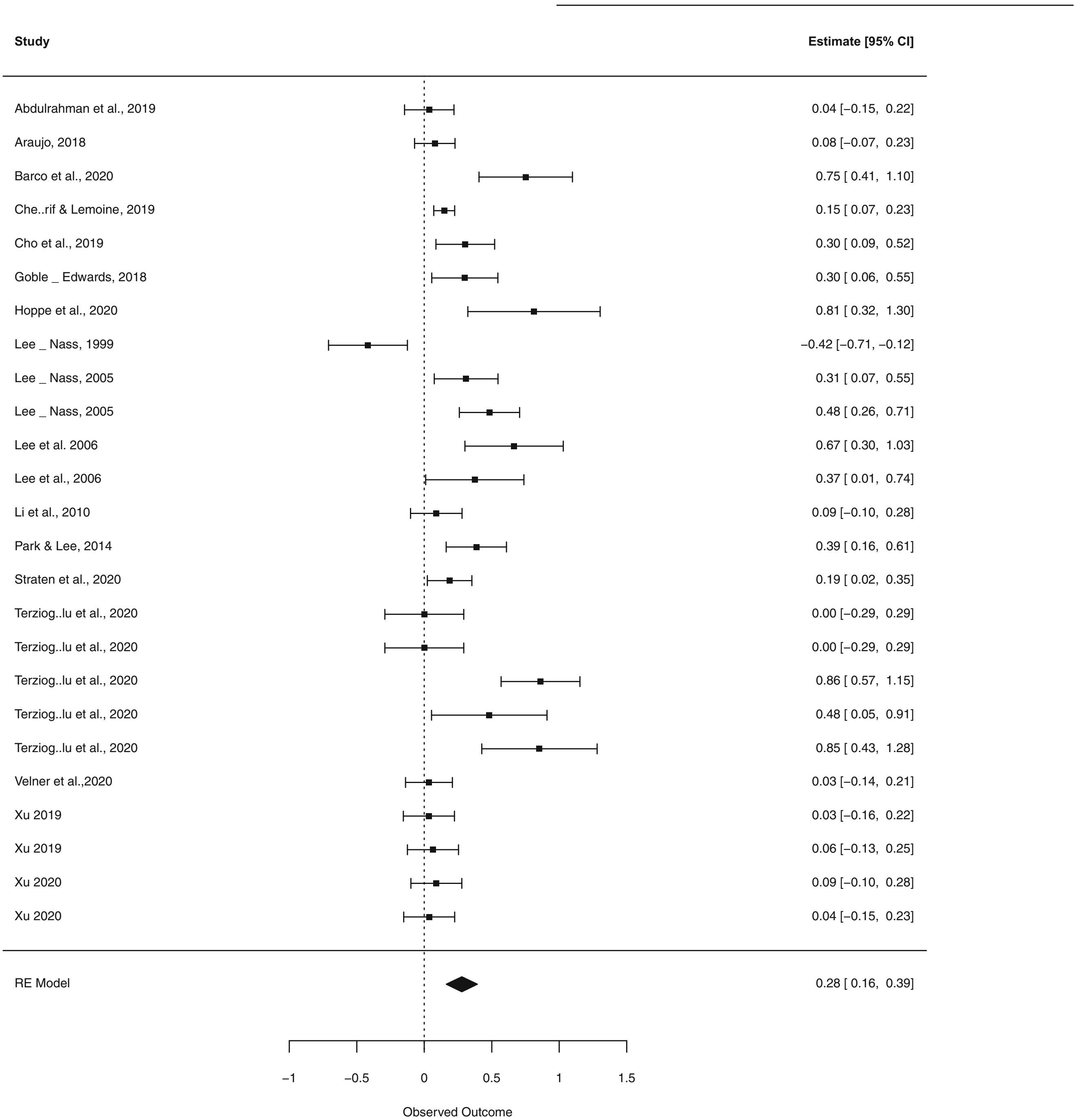
Funnel plot for max-coded meta-analysis



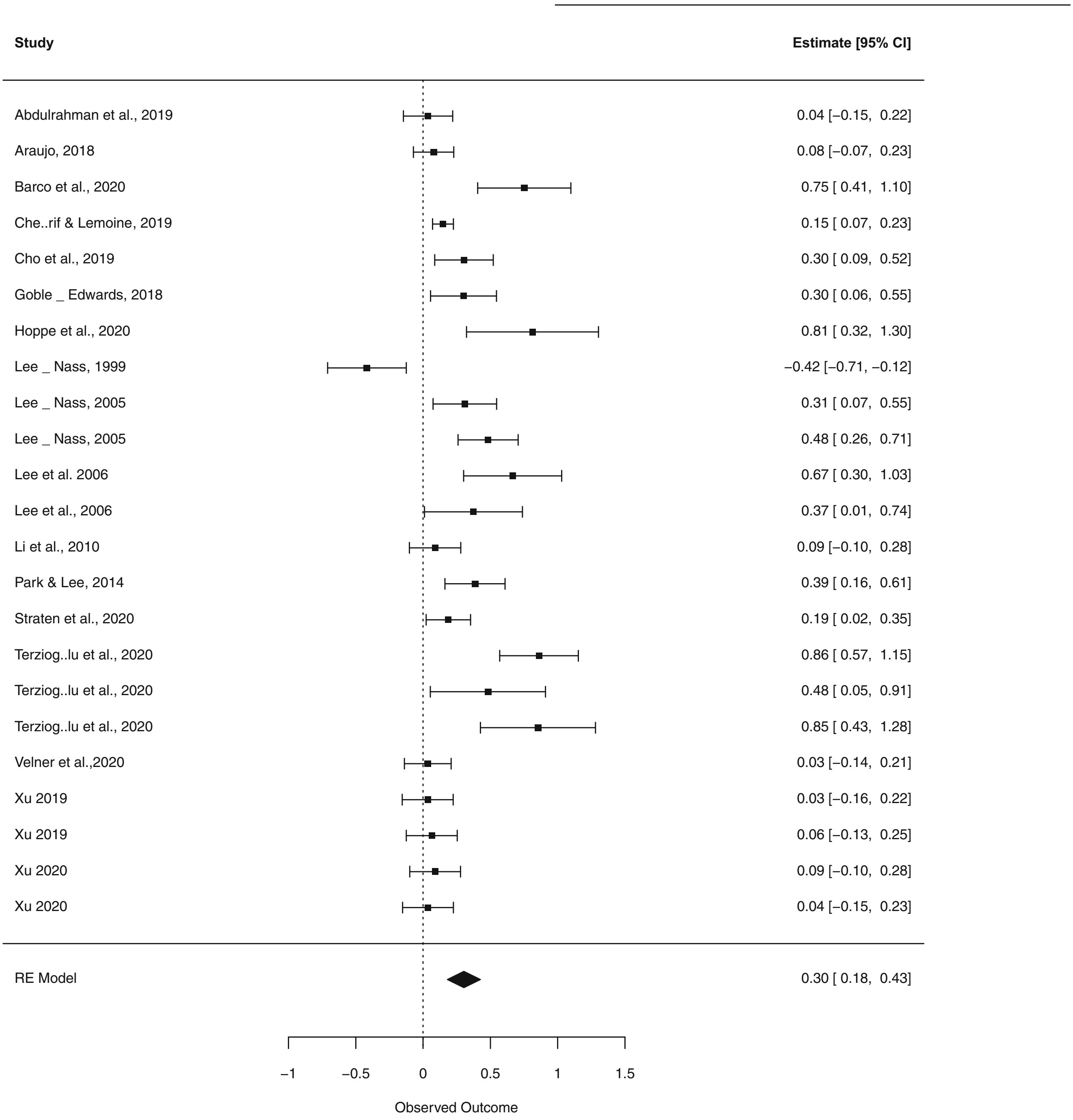
Appendix 3

Forest Plot for Individual Effect Sizes

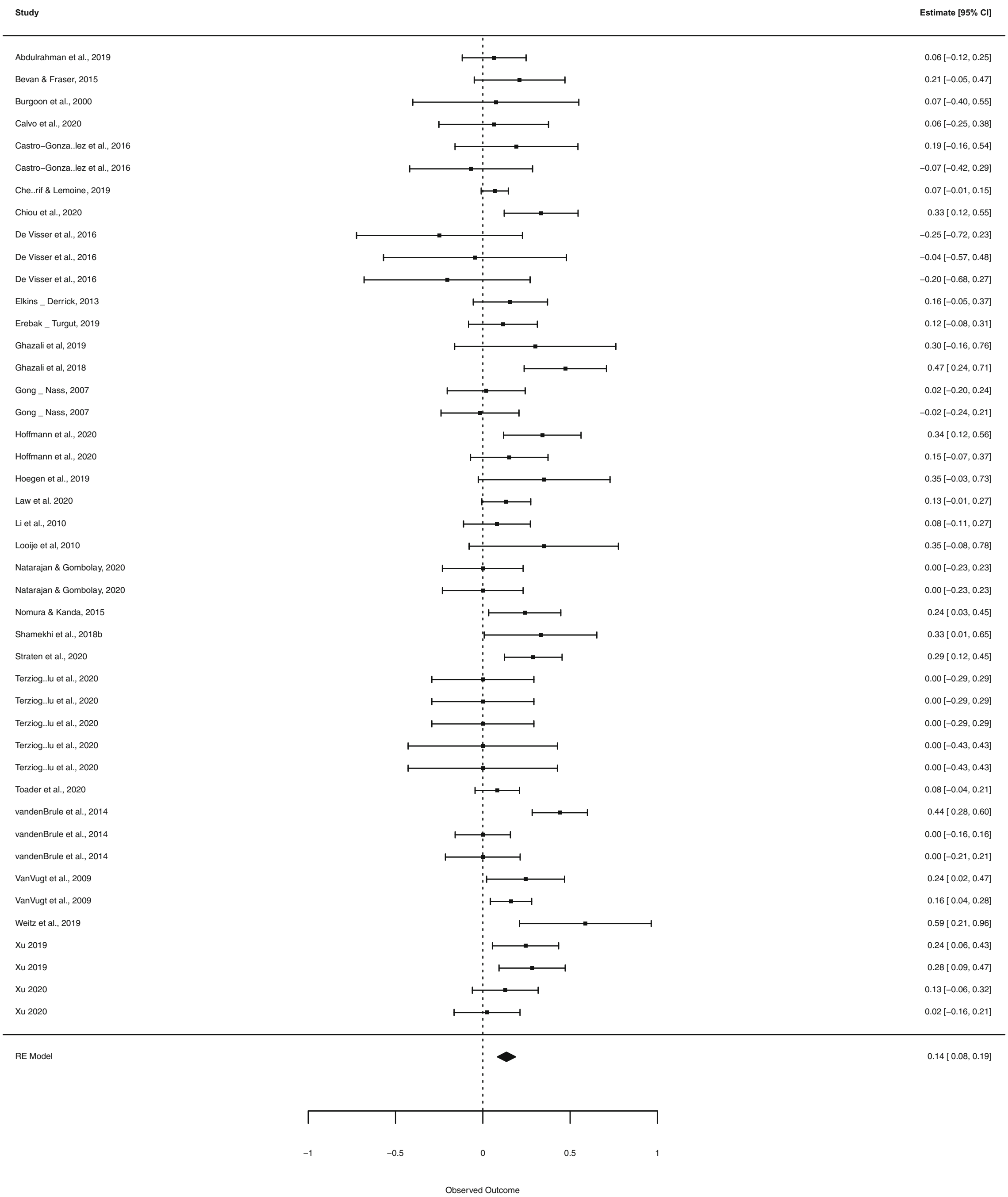
The effects of social cues on social presence (zero-coded)



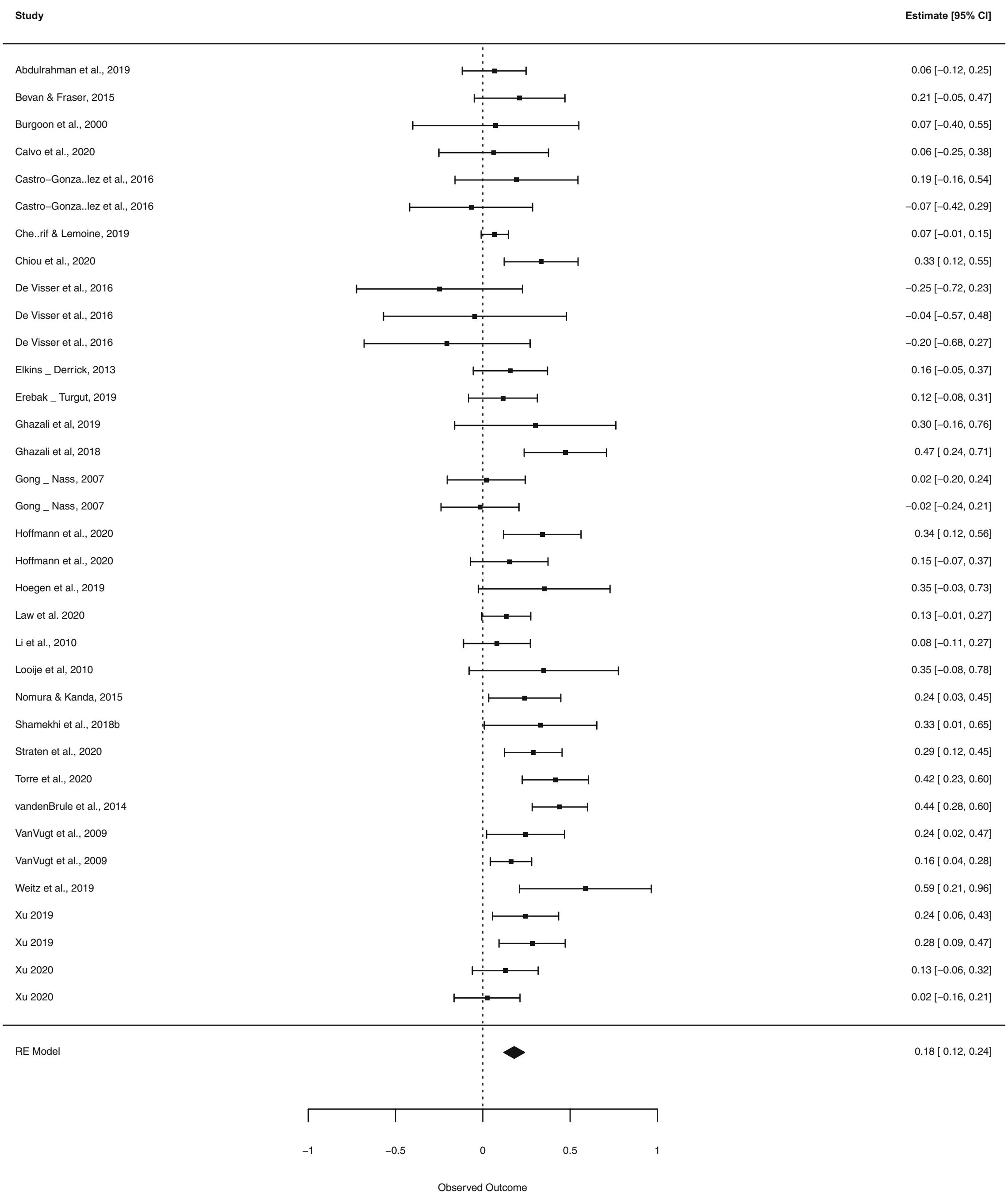
The effects of social cues on social presence (max-coded)



The effects of social cues on trust (zero-coded)



The effects of social cues on trust (max-coded)



References

*means articles included in the meta-analyses

1. *Abdulrahman A, Richards D, Bilgin A (2019) A comparison of human and machine-generated voice. In: Spencer S (ed) Proceedings—VRST 2019: 25th ACM symposium on virtual reality software and technology, vol 41, pp 1–2
2. Abubshait A, Wiese E (2017) You look human, but act like a machine: Agent appearance and behavior modulate different aspects of human–robot Interaction. Front Psychol 8:1393
3. Adkins M, Brashers D (1995) The power of language in computer-mediated groups. Manag Commun Q 8:289–322
4. Andrist S, Mutlu B, Tapus A (2015) Look like me: matching robot personality via gaze to increase motivation. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 3603–3612
5. Araujo T (2018) Living up to the chatbot hype: the influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Comput Hum Behav 85:183–189
6. Banks J (2020) Theory of Mind in social robots: replication of five established human tests. Int J Soc Robot 12:403–414
7. Barco A, de Jong C, Peter J, Kühne R, van Straten C (2020) Robot morphology and children's perception of social robots: an exploratory study. In: 2020 ACM/IEEE international conference on human-robot interaction, pp 125–127
8. Bartneck C, Forlizzi J (2004) A design-centered framework for social human-robot interaction. RO-MAN 2004. In: 13th IEEE international workshop on robot and human interactive communication, pp 591–594
9. *Bevan C, Fraser D (2015) Shaking hands and cooperation in telepresent human-robot negotiation. In: Proceedings of 2015 10th ACM/IEEE international conference on human-robot interaction (HRI), pp 247–254
10. Biocca F, Harms C, Burgoon JK (2003) Toward a more robust theory and measure of social presence: review and suggested criteria. Presence Teleoperators Virtual Environ 12:456–480
11. Borenstein M, Hedges LV, Higgins JP, Rothstein HR (2009) Introduction to meta-analysis. Wiley
12. Breazeal C (2003) Toward sociable robots. Robot Auton Syst 42:167–175
13. *Burgoon JK, Bonito JA, Bengtsson B, Cederberg C, Lundeberg M, Allspach L (2000) Interactivity in human–computer interaction: a study of credibility, understanding, and influence. Comput Hum Behav 16:553–574
14. Calvo N, Elgarf M, Perugia G, Peters C, Castellano G (2020) Can a social robot be persuasive without losing children's trust? In: 2020 ACM/IEEE international conference on human-robot interaction, pp 157–159
15. Carolus A, Binder JF, Muench R, Schmidt C, Schneider F, Buglass SL (2019) Smartphones as digital companions: characterizing the relationship between users and their phones. New Media Soc 21:914–938
16. *Castro-González Á, Admoni H, Scassellati B (2016) Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness, and unpleasantness. Int J Hum Comput Stud 90:27–38
17. Chen Y (2006) Olfactory display: development and application in virtual reality therapy. In: Artificial reality and telexistence—workshops. ICAT 2006. IEEE, pp 580–584
18. *Chérif E, Lemoine JF (2019) Anthropomorphic virtual assistants and the reactions of Internet users: an experiment on the assistant's voice. Rech Appl Mark 34:28–47
19. *Chiou EK, Schroeder NL, Craig SD (2020) How we trust, perceive, and learn from virtual humans: the influence of voice quality. Comput Educ 146:103756
20. *Cho E, Molina MD, Wang J (2019) The effects of modality, device, and task differences on perceived human-likeness of voice-activated virtual assistants. Cyberpsychol Behav Soc Netw 22:515–520
21. Choi S, Liu SQ, Mattila AS (2019) "How may I help you?" Says a robot: examining language styles in the service encounter. Int J Hosp Manag 82:32–38
22. Cohen J (1988) Statistical power analysis for the behavioral sciences. Routledge, New York
23. Cohen J, Cohen P (1983) Applied multiple regression/correlation analysis in behavioral sciences. Erlbaum
24. Cooper H (2015) Research synthesis and meta-analysis: a step-by-step approach. SAGE Publications
25. Craenen B, Deshmukh A, Foster ME, Vinciarelli A (2018) Do we really like robots that match our personality? The case of Big-Five traits, Godspeed scores and robotic gestures. In: 2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN). IEEE, pp 626–631
26. Cummings JJ, Bailenson JN (2016) How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. Media Psychol 19:272–309
27. Duffy BR (2003) Anthropomorphism and the social robot. Robot Auton Syst 42:177–190
28. *Elkins AC, Derrick DC (2013) The sound of trust: voice as a measurement of trust during interactions with embodied conversational agents. Group Decis Negot 22:897–913
29. *Erebak S, Turgut T (2019) Caregivers' attitudes toward potential robot coworkers in elder care. Cogn Technol Work 21:327–336
30. Fiore SM, Wiltshire TJ, Lobato EJC, Jentsch FG, Huang WH, Axelrod B (2013) Toward understanding social cues and signals in human–robot interaction: Effects of robot gaze and proxemic behavior. Front Psychol 4:1–15
31. Fiske ST, Taylor SE (1991) Social cognition, 2nd edn. McGraw Hill
32. Fox J, Ahn SJ, Janssen JH, Yeykelis L, Segovia KY, Bailenson JN (2015) Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. Hum Comput Interact 30:401–432
33. Fox J, Gambino A (2021) Relationship development with humanoid social robots: applying interpersonal theories to human-robot interaction. Cyberpsychol Behav Soc Netw 24:294–299
34. Gambino A, Fox J, Ratan R (2020) Building a stronger CASA: extending the computers are social actors paradigm. Hum Mach Commun 1:71–85
35. Gaudiello I, Zibetti E, Lefort S, Chetouani M, Ivaldi S (2016) Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to iCub answers. Comput Hum Behav 61:633–655
36. Gauthier I, Tarr MJ (1997) Becoming a "Greeble" expert: exploring mechanisms for face recognition. Vis Res 37:1673–1682
37. *Ghazali AS, Ham J, Barakova EI, Markopoulos P (2018) Effects of robot facial characteristics and gender in persuasive human-robot interaction. Front Robot AI 5:73
38. *Ghazali AS, Ham J, Barakova E, Markopoulos P (2019) Assessing the effect of persuasive robots interactive social cues on users' psychological reactance, liking, trusting beliefs and compliance. Adv Robot 33:325–337
39. *Goble H, Edwards C (2018) A robot that communicates with vocal fillers has … Uhhh… greater social presence. Commun Res Rep 35:256–260
40. Gong L (2008) How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Comput Hum Behav 24:1494–1509

41. Gong L, Lai J (2003) To mix or not to mix synthetic speech and human speech? Contrasting impact on judge-rated task performance versus self-rated performance and Attitudinal Responses. Int J Speech Technol 6:123–131
42. *Gong L, Nass C (2007) When a talking-face computer agent is half-human and half-humanoid: human identity and consistency preference. Hum Commun Res 33:163–193
43. Hancock PA, Billings DR, Schaefer KE, Chen JY, De Visser EJ, Parasuraman R (2011) A meta-analysis of factors affecting trust in human-robot interaction. Hum Factors 53:517–527
44. Hanson D, Olney A, Pereira IA, Zielke M (2005) Upending the uncanny valley. AAAI 5:24–31
45. Heeter C (1992) Being there: the subjective experience of presence. Presence Teleoperators Virtual Environ 1:262–271
46. Heider F, Simmel M (1944) An experimental study of apparent behavior. Am J Psychol 57:243–259
47. Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327:557–560
48. Hinds PJ, Roberts TL, Jones H (2004) Whose job is it anyway? A study of human-robot interaction in a collaborative task. Hum Comput Interact 19:151–181
49. Ho A, Hancock J, Miner AS (2018) Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. J Commun 68:712–733
50. *Hoegen R, Aneja D, McDuff D, Czerwinski M (2019) An end-to-end conversational style matching agent. In: Proceedings of the 19th ACM international conference on intelligent virtual agents, pp 111–118
51. *Hoffmann L, Derksen M, Kopp S (2020) What a pity, Pepper! How warmth in robots' language impacts reactions to errors during a collaborative task. In: Companion of the 2020 ACM/IEEE international conference on human-robot interaction, pp 245–247
52. *Hoppe M, Rossmy B, Neumann DP, Streuber S, Schmidt A, Machulla TK (2020) A human touch: social touch increases the perceived human-likeness of agents in virtual reality. In: Proceedings of the 2020 CHI conference on human factors in computing systems, pp 1–11
53. Horstmann AC, Bock N, Linhuber E, Szczuka JM, Straßmann C, Krämer NC (2018) Do a robot's social skills and its objection discourage interactants from switching the robot off? PLoS ONE 13:1–25
54. Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211
55. Johnson BT, Eagly AH (2000) Quantitative synthesis of social psychological research. In: Reis HT, Judd CM (eds) Handbook of research methods in social and personality psychology. Cambridge University Press, pp 496–528
56. Johnson D, Gardner J, Wiles J (2004) Experience as a moderator of the media equation: the impact of flattery and praise. Int J Hum Comput Stud 61:237–258
57. Kim RH, Moon Y, Choi JJ, Kwak SS (2014) The effect of robot appearance types on motivating donation. In: Proceedings of the 2014 ACM/IEEE international conference on human-robot interaction, pp 210–211
58. Kobiella A, Grossmann T, Reid VM, Striano T (2008) The discrimination of angry and fearful facial expressions in 7-month-old infants: an event-related potential study. Cogn Emot 22:134–146
59. Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 4:863
60. Law T, Chita-Tegmark M, Scheutz M (2020) The interplay between emotional intelligence, trust, and gender in human–robot interaction. Int J Soc Robot 13:297–309
61. Lee KM (2004) Presence explicated. Commun Theory 14:27–50
62. Lee EJ (2010) What triggers social responses to flattering computers? Experimental tests of anthropomorphism and mindlessness explanations. Commun Res 37:191–214
63. *Lee KM, Jung Y, Kim J, Kim SR (2006) Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. Int J Hum Comput Stud 64:962–973
64. *Lee EJ, Nass C (1999) Effects of the form of representation and number of computer agents on conformity. In: CHI '99 extended abstracts on human factors in computing systems, pp 238–239
65. *Lee KM, Nass C (2005) Social psychological origins of feelings of presence: creating social presence with machine-generated voices. Media Psychol 7:31–45
66. Lee KM, Peng W, Jin SA, Yan C (2006) Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. J Commun 56:754–772
67. Leichtmann B, Nitsch V (2020) How much distance do humans keep toward robots? Literature review, meta-analysis, and theoretical considerations on personal space in human-robot interaction. J Environ Psychol 68:101386
68. Li J (2015) The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum Comput Stud 77:23–37
69. Li J, Chignell M (2011) Communication of emotion in social robots through simple head and arm movements. Int J Soc Robot 3:125–142
70. Li JJ, Ju W, Reeves B (2017) Touching a mechanical body: tactile contact with body parts of a humanoid robot is physiologically arousing. J Hum Robot Interact 6:118–130
71. Li J, Kizilcec R, Bailenson J, Ju W (2016) Social robots and virtual agents as lecturers for video instruction. Comput Hum Behav 55:1222–1230
72. Li D, Rau PLP, Li Y (2010) A cross-cultural study: effect of robot appearance and task. Int J Soc Robot 2:175–186
73. Lombard M, Ditton T (1997) At the heart of it all: the concept of presence. J Comput Mediat Commun 3(2)
74. Lombard M, Xu K (2021) Social responses to media technologies: the Media are Social Actors paradigm. Hum Mach Commun 2:29–55
75. *Looije R, Neerincx MA, Cnossen F (2010) Persuasive robotic assistant for health self-management of older adults: design and evaluation of social behaviors. Int J Hum Comput Stud 68:386–397
76. Martin D, Macrae CN (2007) A face with a cue: exploring the inevitability of person categorization. Eur J Soc Psychol 37:806–816
77. Martini MC, Gonzalez CA, Wiese E (2016) Seeing minds in others: Can agents with robotic appearance have human-like preferences? PLoS ONE 11:e0146310
78. Mayer RE, Sobko K, Mautone P (2003) Social cues in multimedia learning: role of speaker's voice. J Educ Psychol 95:419–425
79. Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLOS Med 6:e1000097
80. Morewedge CK, Preston J, Wegner DM (2007) Timescale bias in the attribution of mind. J Pers Soc Psychol 93:1–11
81. Mori M, MacDorman KF, Kageki N (2012) The uncanny valley. IEEE Robot Autom Mag 19:98–100
82. Naneva S, Gou MS, Webb TL, Prescott TJ (2020) A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. Int J Soc Robot 12:1179–1201
83. Nass C (2004) Etiquette equality: Exhibitions and expectations of computer politeness. Commun ACM 47:35–37

84. Nass C, Brave S (2005) Wired for speech: how voice activates and advances the human-computer relationship. MIT Press, Cambridge
85. Nass C, Fogg BJ, Moon Y (1996) Can computers be teammates? Int J Hum Comput Stud 45:669–678
86. Nass C, Lombard M, Henriksen L, Steuer J (1995) Anthropocentrism and computers. Behav Inf Technol 14:229–238
87. Nass C, Moon Y (2000) Machines and mindlessness: social responses to computers. J Soc Issues 56:81–103
88. Nass C, Moon Y, Fogg BJ, Reeves B, Dryer C (1995) Can computer personalities be human personalities? In: Proceedings of conference companion on human factors in computing systems, pp 228–229
89. Nass C, Moon Y, Green N (1997) Are machines gender neutral? Gender-stereotypic responses to computers with voices. J Appl Soc Psychol 27:864–876
90. Nass C, Reeves B, Leshner G (1996) Technology and roles: a tale of two TVs. J Commun 46:121–128
91. Nass C, Steuer J, Tauber ER (1994) Computers are social actors. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 72–78
92. *Natarajan M, Gombolay M (2020) Effects of anthropomorphism and accountability on trust in human robot interaction. In: Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction, pp 33–42
93. *Nomura T, Kanda T (2015) Influences of evaluation and gaze from a robot and humans' fear of negative evaluation on their preferences of the robot. Int J Soc Robot 7:155–164
94. Nomura T, Kanda T, Suzuki T, Kato K (2009) Age differences and images of robots: social survey in Japan. Interact Stud 10:374–391
95. Nomura T, Yamada S, Kanda T, Suzuki T, Kato K (2009) Influences of concerns toward emotional interaction into social acceptability of robots. In: 2009 4th ACM/IEEE international conference on human-robot interaction (HRI), pp 231–232
96. Oh CS, Bailenson JN, Welch GF (2018) A systematic review of social presence: definition, antecedents, and implications. Front Robot AI 5:114
97. Okumura Y, Kanakogi Y, Kanda T, Ishiguro H, Itakura S (2013) Infants understand the referential nature of human gaze but not robot gaze. J Exp Child Psychol 116:86–95
98. *Park E, Lee J (2014) I am a warm robot: the effects of temperature in physical human-robot interaction. Robotica 32:133–142
99. Perez S (2020). Duplex, Google's conversational AI, has updated 3M+ business listings since pandemic. https://techcrunch.com/2020/10/15/duplex-googles-conversational-a-i-has-updated-3m-business-listings-since-pandemic/. Accessed 15 Oct 2020
100. Pfeifer R, Scheier C (1999) Understanding intelligence. MIT Press, Cambridge
101. Rains SA, Matthes J, Palomares NA (2020) Communication science and meta-analysis: introduction to the special issue. Hum Commun Res 46:115–119
102. Reeves B, Nass C (2000) Perceptual user interfaces: perceptual bandwidth. Commun ACM 43:65–70
103. Reeves B, Nass C (2002) The media equation: How people treat computers, television, and new media like real people and places. CSLI Publications
104. Rosenthal R (1979) The file drawer problem and tolerance for null results. Psychol Bull 86(3):638–641
105. Rosenthal R (1991) Meta-analytic procedures for social research. SAGE, Newbury Park
106. Rosenthal R (1995) Writing meta-analytic reviews. Psychol Bull 118:183–192
107. Rosenthal R, DiMatteo MR (2001) Meta-analysis: recent developments in quantitative methods for literature reviews. Annu Rev Psychol 52:59–82
108. Sah YJ, Peng W (2015) Effects of visual and linguistic anthropomorphic cues on social perception, self-awareness, and information disclosure in a health website. Comput Hum Behav 45:392–401
109. Salem M, Eyssel F, Rohlfing K, Kopp S, Joublin F (2013) To err is human(-like): effects of robot gesture on perceived anthropomorphism and likability. Int J Soc Robot 5:313–323
110. Santamaria T, Nathan-Roberts D (2017) Personality measurement and design in human-robot interaction: a systematic and critical review. In: Proceedings of the human factors and ergonomics society annual meeting, vol 61, pp 853–857
111. Schmidt KL, Cohn JF (2001) Human facial expressions as adaptations: evolutionary questions in facial expression research. Am J Phys Anthropol 116:3–24
112. Serrano JM, Iglesias J, Loeches A (1992) Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4-to 6-month-old infants. Dev Psychobiol J Int Soc Dev Psychobiol 25:411–425
113. *Shamekhi A, Liao QV, Wang D, Bellamy RK, Erickson T (2018) Face Value? Exploring the effects of embodiment for a group facilitation agent. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–13
114. Sherry JL (2001) The effects of violent video games on aggression: a meta-analysis. Hum Commun Res 27:409–431
115. Shin DH, Choo H (2011) Modeling the acceptance of socially interactive robotics: Social presence in human–robot interaction. Interact Stud 12:430–460
116. Spears R, Postmes T (2015) Group identity, social influence, and collective action online. Extensions and applications of the SIDE model. In: Sundar S (ed) The handbook of the psychology of communication technology. Wiley, New York, pp 23–46
117. Stock-Homburg R, Hannig M, Lilienthal L (2020) Conversational flow in human-robot interactions at the workplace: comparing humanoid and android robots. In: Wagner AR et al (eds) Social Robotics. ICSR 2020. Lecture notes in computer science, vol 12483. Springer, Cham. https://doi.org/10.1007/978-3-030-62056-1_4
118. Stower R, Calvo-Barajas N, Castellano G, Kappas A (2021) A meta-analysis on children's trust in social robots. Int J Soc Robot 13:1979–2001
119. *Straten CLV, Peter J, Kühne R, Barco A (2020) Transparency about a robot's lack of human psychological capacities: effects on child-robot perception and relationship formation. ACM Trans Hum Robot Interact THRI 9:1–22
120. Sundar SS (2020) Rise of machine agency: a framework for studying the psychology of human–AI interaction (HAII). J Comput Mediat Commun 25(1):74–88
121. *Terzioğlu Y, Mutlu B, Şahin E (2020) Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration. In: Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction, pp 343–357
122. *Torre I, Goslin J, White L (2020) If your device could smile: people trust happy-sounding artificial agents more. Comput Hum Behav 105:106215
123. Torre I, Goslin J, White L, Zanatto D (2018) Trust in artificial voices: a "congruency effect" of first impressions and behavioural experience. In: Proceedings of the technology, mind, and society, pp 1–6
124. Treal T, Jackson PL, Meugnot A (2020) Combining trunk movement and facial expression enhances the perceived intensity and believability of an avatar's pain expression. Comput Hum Behav 112:106451
125. Tung FW, Deng YS (2007) Increasing social presence of social actors in e-learning environments: effects of dynamic and static emoticons on children. Displays 28:174–180

126. *van den Brule R, Dotsch R, Bijlstra G, Wigboldus DH, Haselager P (2014) Do robot performance and behavioral style affect human trust? Int J Soc Robot 6:519–531
127. *van Vugt HC, Konijn EA, Hoorn JF, Veldhuis J (2009) When too heavy is just fine: creating trustworthy e-health advisors. Int J Hum Comput Stud 67:571–583
128. *Velner E, Boersma PP, de Graaf MM (2020) Intonation in robot speech: does it work the same as with people? In: Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction, pp 569–578
129. *de Visser EJ, Monfort SS, McKendrick R, Smith MAB, McKnight PE, Krueger F, Parasuraman R (2016) Almost human: anthropomorphism increases trust resilience in cognitive agents. J Exp Psychol Appl 22:331–349
130. Walters ML, Lohse M, Hanheide M, Wrede B, Syrdal DS, Severinson-Eklundh K (2011) Evaluating the robot personality and verbal behavior of domestic robots using video-based studies. Adv Robot 25:2233–2254
131. Wang LC, Baker J, Wagner JA, Wakefield K (2007) Can a retail web site be social? J Mark 71:143–157
132. Wang B, Rau PLP (2019) Influence of embodiment and substrate of social robots on users' decision-making and attitude. Int J Soc Robot 11:411–421
133. *Weitz K, Schiller D, Schlagowski R, Huber T, André E (2019) "Do you trust me?" Increasing user-trust by integrating virtual agents in explainable AI interaction design. In: Proceedings of the 19th ACM international conference on intelligent virtual agents, pp 7–9
134. Westerman D, Cross AC, Lindmark PG (2019) I believe in a thing called bot: perceptions of the humanness of "chatbots." Commun Stud 70:295–312
135. Woods S, Dautenhahn K, Kaouri C, te Boekhorst R, Koay KL, Walters ML (2007) Are robots like people? Relationships between participant and robot personality traits in human–robot interaction studies. Interact Stud 8:281–305
136. *Xu K (2019) First encounter with robot Alpha: How individual differences interact with vocal and kinetic cues in users' social responses. New Media Soc 21:2522–2547
137. *Xu K (2020) Language, modality, and mobile media use experiences: social responses to smartphone cues in a task-oriented context. Telemat Inform 48:101344
138. You S, Robert LP (2018) Human-robot similarity and willingness to work with a robotic co-worker. In: Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction, pp 251–260
139. Zhao S (2003) Toward a taxonomy of copresence. Presence Teleoperators Virtual Environ 12:445–455
140. Zhao S (2006) Humanoid social robots as a medium of communication. New Media Soc 8:401–419
141. Ziemke T (2003) What's the thing called embodiment? In: Proceedings of the annual meeting of the cognitive science society, vol 25, pp 1305–1310