Current Status and Future Prospects of Human-Robot Interaction Research: A Preliminary Discussion on the Intersection of Psychology and Artificial Intelligence

Authors: Deng Zhaoxin, Fu Chao, Yang Xue, Wang Yiwen, Wang Yiwen

Date: 2018-12-26T00:00:00+00:00

Abstract

Robots are progressively integrating into our society. While prior research has predominantly focused on robotic software programming or mechanical manufacturing, scant attention has been devoted to the significance of psychology in human-robot interaction studies, particularly concerning the establishment, development, and evolution of human-robot dyadic relationships. Given that social robots possess higher sociability than industrial robots, the term "robot" in human-robot interaction research generally refers specifically to social robots. Building upon an introduction to the principal applications of social robots, this article categorically summarizes research themes in human-robot interaction, including trust, empathy, and social distance; elaborates on the uncanny valley, media equation, and mind perception theories involved in such research; discusses potential ethical issues arising from social robots; and offers targeted prospects for the deepening and expansion of research themes like trust and empathy in future human-robot interaction studies. Psychology can provide knowledge regarding human emotion, cognition, and behavior, thereby facilitating the establishment and shaping of robust human-robot dyadic relationships.

Full Text

Preamble

Deng Zhaoxin[1,2], Fu Chao[1,2], Yang Xue[1,2], Wang Yiwen[1,2]

[1]Institute of Psychological and Cognitive Sciences, Fuzhou University, Fuzhou 350116, China
[2]Center for China Social Trust Studies, Fuzhou University, Fuzhou 350116, China

Abstract

Robots are gradually being integrated into our society. While previous research has primarily focused on software programming or mechanical manufacturing of robots, few studies have highlighted the importance of psychology in human-robot interaction research, particularly in the establishment, development, and transformation of the binary relationship between humans and robots. Given that social robots possess greater social capabilities than industrial robots, the term "robot" in human-robot interaction research generally refers specifically to social robots. Building upon an introduction to the main applications of social robots, this paper categorically summarizes research themes in human-robot interaction, including trust, empathy, and social distance. It elaborates on the theories of the uncanny valley, media equation, and mind perception involved in these studies, discusses potential ethical issues arising from social robots, and offers targeted prospects for the deepening and expansion of future research themes such as trust and empathy in human-robot interaction. Psychology can provide knowledge about human emotion, cognition, and behavior, thereby helping to establish and shape a sound binary relationship between humans and robots.

Keywords: human-robot interaction; social robots; artificial intelligence psychology

1. The Origin of Human-Robot Interaction Research

Science fiction films often depict robots with human-like appearance and thoughts, capable of experiencing emotions, expressing ideas, and taking actions just like humans. How far is such a scenario from real life? Will future robots truly be able to think and feel like humans? The history of robot manufacturing can be traced back to 1459, when Leonardo da Vinci designed a mechanical knight that could move its arms and raise its visor through pulley systems (Rosheim, 2006). In the mid-20th century, Turing (1950) laid a solid foundation for the development of autonomous robots—machines that can complete tasks independently without continuous human guidance. In the 1950s, the first generation of true robots emerged as programmable mechanical arms for the automotive industry. Since then, industrial robots have continuously evolved and been gradually applied in military domains such as surveillance, bomb disposal, and automated weapons. However, whether industrial or military, these robots were mostly regarded as tools rather than work partners or comrades, and no general sense of interaction was formed between humans and these machines. In psychology, interaction occurs between two or more social entities that exert mutual influence and effect on each other. Therefore, the prerequisite for human-robot interaction is that robots are perceived as partners rather than tools. How people perceive robots depends largely on their functions, which can be divided into practical and social categories. Practical functions involve liberating human labor and improving production efficiency, whereas social functions involve providing social support and companionship. Traditionally, robots have primarily served practical functions, but with technological advances over the past two decades, their functions have gradually shifted toward social capabilities (Wang & Krumhuber, 2018). Robots have slowly integrated into daily life, becoming capable assistants in homes, schools, hospitals, and other environments. These social robots no longer serve merely as tools but as intimate human partners. This shift has raised questions such as: What factors influence the efficiency of human-robot collaboration, and do humans bear responsibility toward robots? These questions have gradually given rise to human-robot interaction (HRI) research. Given the high social nature of social robots, this paper focuses specifically on interaction between humans and social robots. The following sections provide a brief overview of social robot applications and elaborate on the current state and future directions of human-robot interaction research.

2. Main Applications of Social Robots

Social robots are gradually entering people' s lives. In some nursing homes, medical robots provide daily care for the elderly; in certain educational institutions, teaching robots assist with children' s education; and in large shopping centers, guide robots offer directional information. Researchers strive to manufacture robots with diverse forms and functions to meet the needs of different user groups. This section briefly reviews the current applications of social robots.

2.1 Medical Robots

Medical robots primarily serve two groups: the elderly and children. The world is facing a population aging crisis, particularly in developed countries where healthcare workers cannot meet the growing medical needs of the elderly population. Increasingly, people are turning to robotic assistance to address this challenge (Mann, MacDonald, Kuo, Li, & Broadbent, 2015). Robots can help monitor blood pressure, provide reminders for forgetful individuals, and assist those with mobility impairments in walking and bathing. Moreover, in terms of mental health, robots can offer companionship and communication. The most famous companion social robot is Paro from Japan, which resembles a Greenland seal pup (see Figure 1 [Figure 1: see original paper]-a). When sensors detect touch, light, sound, or position changes, Paro emits seal-like vocalizations. Research indicates that the appearance of robots caring for the elderly should match their tasks—people prefer furry robots for companionship over mechanical-looking ones (Broadbent, Tamagawa, Kerse, & Knock, 2009). Therefore, designers chose the seal pup image, hoping people would hug, pet, and communicate with Paro as they would with a pet dog. Paro is primarily used in nursing homes, where it provides companionship to the elderly and plays a role in treating dementia. A randomized controlled trial with elderly participants in nursing homes showed that, compared to real pet dogs, Paro significantly reduced loneliness and increased social interaction behaviors, such as more communication with the robot or others (Robinson, Macdonald, Kerse,

& Broadbent, 2013).

Regarding children, research indicates that children' s behavioral patterns differ from adults' . During interaction with robots, children tend to view robots as living entities, giving child-robot interaction a special status in HRI research (Vallèsperis, Angulo, & Domènech, 2018). This study analyzed children' s imaginings of interacting with social robots in medical contexts to explore the characteristics medical robots should possess. In practical applications, medical social robots can induce children to perform certain behaviors to assess their level of autism for diagnosis and treatment, while also teaching life skills and increasing prosocial behaviors (Diehl, Schmitt, Villano, & Crowell, 2012). Children with autism spectrum disorders have difficulty understanding others' intentions and communicating with them. Compared to humans, robots' states are more predictable and their range of activities is more limited, making interaction easier for these children. Through interaction with robots, children with autism can learn social norms and apply them to real human interactions. However, due to large individual differences among subjects, insufficient sample sizes, and lack of proper control groups, the effectiveness of robot-assisted therapy for children with autism needs further improvement (Diehl et al., 2012). Future research requires the involvement of clinicians and psychologists to enhance the effectiveness of such robot-assisted therapies through better experimental designs.

2.2 Educational Robots

Educational robots in teaching fall into two categories: first, robots that can be freely assembled and support visual programming (such as LEGO Mindstorms, see Figure 1-b), which can serve as teaching aids or help teachers impart knowledge; and second, interactive humanoid robots that can directly act as teachers to instruct students in foreign languages and other subjects (Mubin, Stevens, Shahid, Mahmud, & Dong, 2013). One study examined the difference in teaching effectiveness between robot and human teachers in instructing children in foreign languages. The results showed that children were more proactive in communicating in the foreign language with robot teachers, whereas they appeared shy and hesitant with human teachers. Additionally, robots can work continuously, overcoming the fatigue problems that human teachers may experience (Chang, Lee, Chao, Wang, & Chen, 2010). Another study investigated the teaching effectiveness of such robots for children from low-income families, finding that robot teachers could enhance children' s motivation more effectively than human teachers, while also strengthening their sense of community and self-expression (Han, Park, & Park, 2015).

2.3 Guide Robots

In large public places such as shopping malls and museums, people often encounter guide robots. Besides attracting customers, guide robots can provide venue information and distribute flyers. Sabelli and Kanda (2015) conducted interviews with customers in a Japanese shopping mall equipped with guide robots, revealing that people viewed the robot more as a mascot than a public facility, and customers appreciated the robot' s presence and supported its placement as a guide. Another interview study showed that most customers considered guide robots providing directions and distributing flyers to be beneficial. Guide robots do not "judge people by their appearance" as humans might—65% of respondents preferred robots over humans for the same services, and over 90% expressed willingness to use guide robots again (Satake, Hayashi, Nakatani, & Kanda, 2015).

3. Research Themes in Human-Robot Interaction

Human-like robots and pet-like robots are the two main categories of robots studied in human-robot interaction research. Human-like robots are those that closely resemble humans in both appearance and behavior, designed to enable more natural human-robot interaction through instinctual responses. "Jiajia," the third-generation unique experience interactive robot developed by the University of Science and Technology of China, is China' s first human-like social robot. Born in April 2016, standing 1.6 meters tall, it features delicate facial features and skin texture close to that of real humans (see Figure 1-c), with capabilities including human-computer dialogue comprehension, facial micro-expression display, and coordinated mouth and body movements. The research team believes that a robot' s image should be consistent with its character and functions. They were the first to propose and explore the definition of robot character, endowing "Jiajia" with traits such as kindness, diligence, and wisdom. Abroad, research on highly realistic human-like robots also attracts scientists. Japanese roboticist Ishiguro built androids (see Figure 1-d) that are exact replicas of real people in appearance. Through remote control, these robots can exhibit behaviors, conversational abilities, and even personalities consistent with the actual person. Researchers believe such robots can reflect specific human images and help researchers gain deeper understanding of human nature (Ishiguro & Nishio, 2007). Additionally, Phillips' research team at Brown University studied the degree of anthropomorphism in human-like robots. They argued that the 外观设计 of human-like robots mostly relies on developers' intuition and lacks systematic understanding of the relationships among the scope, types, and constituent features of human-like robots. Accordingly, the team developed an online assessment tool to help researchers evaluate the anthropomorphic degree of human-like robots (Phillips, Zhao, Ullman, & Malle, 2018).

Pet-like robots are robots whose appearance resembles small animals or pets. The Paro developed by Japan' s National Institute of Advanced Industrial Science and Technology and Pleo developed by UGOBE mentioned earlier are typical pet-like robots. Pleo has an appearance resembling a dinosaur hatchling (see Figure 1-e), can learn voice commands, and its built-in sensors can perceive feeding and respond to human touch. Some design principles have been integrated into existing pet-like robots, such as robots having the freedom to reject or accept commands, owners helping robots grow and taking responsibility for them during this process, and enabling robots to develop dependence on their owners (Kaplan, 2001). Research on pet-like robot interaction can analyze the pet behavior patterns people need, allowing researchers to design and manufacture more acceptable pet-like robots.

[Figure 2: see original paper]

Overall, human-robot interaction research involves multiple disciplines. Engineers are responsible for manufacturing social robots, while psychologists use these robots to interact with real people to explore factors influencing the human-robot binary relationship and feed these research findings back to engineers, enabling them to further develop robots with stronger interaction capabilities (Broadbent, 2017), thus forming the research loop shown in Figure 2. Throughout this interdisciplinary research process, the importance of psychological factors has gradually increased. Pursuing highly complex functionality is not the primary goal of human-robot interaction research. A feasible research strategy is to first consider the specific needs of robot users and then develop robot functions based on those needs (Vincent, Taipale, Sapio, Lugano, & Fortunati, 2015). In other words, feedback from psychologists in the research loop is crucial. This human-centered robot design concept has led many researchers to explore psychological issues involved in human-robot interaction. The following sections categorize and review existing research themes.

3.1 Trust in Human-Robot Interaction

Human trust in robots has gradually attracted researchers' attention as robots have developed. This trust significantly influences the effectiveness of human-robot collaboration. At high trust levels, people may over-rely on and misuse robots, while at low trust levels, they may completely abandon them. Examining human trust levels in robots is thus an important research theme.

Currently, four main paradigms are used to study trust issues in human-robot interaction:

(1) **Interviews**: In a study exploring social robot acceptance, researchers used semi-structured interviews to assess people's trust in the robot Nabaztag (see Figure 1-f). Nabaztag was brought into participants' homes to provide weather forecasts and exercise recommendations. Through interview content analysis, researchers found that most participants trusted the robot and its information, but trust levels decreased when the robot provided incorrect information (de Graaf, Allouch, & Klamer, 2015).

(2) **Modeling**: One study simulated battlefield environments to establish an evaluation system for human-robot collaborative teams, investigating human trust in tactical robots during collaborative tasks (Freedy, Devisser, Weltman, & Coeyman, 2008). The research team developed a multidimensional collaborative performance model covering various factors influencing human-robot collaborative trust (such as task allocation and robot

controllability) and evaluated the importance of these factors through simulation experiments.

(3) **Trust Game**: In laboratory settings, researchers often use trust game tasks to examine human trust in robots. The general form of the trust game (Berg, Dickhaut, & McCabe, 1995) involves two anonymous players who each possess a certain amount of money S. One player (the trustor) is asked to give a portion of money Y (0 Y S) to the other player (the trustee). The trustee then receives 3Y and decides how much money X (0 X 3Y) to return to the trustor. The trustor' s final payoff is S - Y + X, and the trustee' s payoff is S + 3Y - X. In human-robot trust games, participants are typically asked to act as trustors and complete the task with a robot. Using this paradigm, Haring et al. (2013) analyzed human trust in human-like robots and found that more extroverted participants were more inclined to give money to the robot trustee, indicating that more extroverted individuals had higher trust in robots.

(4) **Prisoner' s Dilemma Game (PDG)**: In addition to trust games, some researchers have used the prisoner' s dilemma game to assess trust and reciprocity in human-robot interaction. In this game, two prisoners face a choice: if both choose to cooperate (remain silent), both receive short sentences; if one chooses to betray, the betrayer is released while the betrayed receives a long sentence. One study used repeated prisoner' s dilemma games to examine trust reciprocity in human-robot interaction (Sandoval, Brandstetter, Obaid, & Bartneck, 2015). In the experiment, reciprocity was defined as giving the same type of feedback in response to friendly/hostile behavior—after a robot or human agent made a cooperative/betrayal choice, the participant made the same choice. The results showed that participants were more inclined to cooperate with humans than with robots, indicating a greater tendency to trust human agents.

3.2 Empathy in Human-Robot Interaction

Empathy refers to the ability to understand others' emotions and feelings by putting oneself in their shoes (Decety & Jackson, 2004). It is generally considered an important capability in social interaction and a foundation for social cooperation and prosocial behavior. Empathy between humans and robots is a key factor influencing the effectiveness of human-robot collaboration. In laboratory research, empathy between humans and robots is typically studied through scenario priming paradigms. In one study, Leite et al. (2013) had the robot iCat (see Figure 1-g) observe a chess game between two human players and comment on every move. The experiment set iCat to make empathetic comments toward one player and neutral comments toward the other to examine the impact of empathetic versus neutral feedback on human-robot empathy. Analysis of open-ended questionnaires completed by participants after the experiment revealed that those treated with empathy by the robot perceived it as friendlier. Other researchers have primed empathy by having participants watch videos of the robot pet Pleo and real humans being caressed or abused, using functional magnetic resonance imaging (fMRI) combined with scales to investigate this phenomenon. The results showed that when participants watched videos of either robots or humans being caressed, the patterns of brain activation were identical. However, participants showed greater distress when watching videos of human abuse compared to robot abuse (Rosenthal-von der Pütten et al., 2014). Another study primed empathy by presenting participants with images of robotic hands being cut by scissors and used electroencephalography (EEG) to examine differences in empathy toward humans and robots. The results indicated that differences in empathy toward humans and robots mainly occurred in the early stages of top-down processing, while the later stages were similar (Suzuki, Galli, Ikeda, Itakura, & Kitazaki, 2015).

3.3 Social Distance in Human-Robot Interaction

In interpersonal communication, people typically exhibit different social distances based on their degree of closeness or remoteness, manifested in both physical and psychological aspects. Does this phenomenon also exist in human-robot interaction? Sociologist Park (1942) proposed the concept of social distance, describing it as the degree of lack of intimacy between people in terms of ethnicity, race, religion, occupation, and other aspects. One study divided social distance in human-robot interaction into three dimensions (power distance, task distance, and spatial distance) and examined different human responses to robots when social distance varied (Kim & Mutlu, 2014). Power distance refers to the superior-subordinate status relationship between humans and robots, task distance refers to the structure of task execution (cooperation or competition), and spatial distance refers to the physical proximity between humans and robots. Participants completed a card-matching task with robots under different social distance conditions. The results showed that when participants faced subordinate robots, a greater spatial distance provided a better experience. Additionally, participants rated their experience of cooperating with robots more positively than competing against them. These results emphasize the importance of consistency between social distance and behavioral performance in human-robot interaction, suggesting that future research should further explore how social distance factors influence human-robot relationships.

3.4 Preference for Robots

Thanks to the convenience brought by robots, humans show preferences for robots over other devices or even real humans in certain aspects. One study compared human evaluations of robots and tablets (Mann et al., 2015). Researchers had the robot iRobiQ (see Figure 1-h) and a tablet ask participants health-related questions and provide identical exercise and relaxation recommendations. The results showed that participants interacted more positively with the robot (evidenced by increased verbal responses and smiles) and were more willing to accept recommendations from the robot. Interview analyses also indicated that participants evaluated the robot and human-robot relationships more positively and expressed greater willingness to interact with the robot again. Another study revealed a human preference for robots over humans in guidance services (Satake et al., 2015). Researchers followed up with customers who used guide robots in a shopping mall and found that, compared to human guides, customers preferred robot guides for services such as distributing flyers and providing directions. This preference was not due to superior service quality but rather to curiosity about robots. Additionally, customers did not feel the embarrassment or shyness when requesting services from robots that they might experience with human staff.

Overall, in these research themes on human-robot interaction, researchers attempt to explore the establishment and development of the human-robot binary relationship from a psychological perspective. Human-robot interaction research requires support from relevant theories, which are briefly summarized below.

4.1 Uncanny Valley Theory

Mori (1970) was one of the first scientists to study human-robot interaction. In the 1970s, he proposed that the relationship between a robot' s degree of anthropomorphism and human comfort level changes as shown in Figure 3 [Figure 3: see original paper]: the more human-like a robot' s appearance, the higher the comfort level, but when a robot' s appearance becomes extremely close to yet not completely identical to a real human, comfort drops sharply. This phenomenon is called the uncanny valley (Mori, 1970; 邓卫斌, 于国龙, 2016). Mori also explained the effect of movement on the uncanny valley, noting that a robot's movement patterns also alter human comfort levels. To make people feel natural and comfortable when interacting with robots, human-like social robots should have as high a degree of anthropomorphism as possible while avoiding falling into the "uncanny valley" due to excessive human-likeness. Some researchers have explained the uncanny valley theory: when participants observe robot behaviors that do not match expectations, error feedback signals are generated in the brain. In other words, if a robot looks human-like but moves in a strange manner—when it resembles humans in "form" but not in "action"—a conflict with intuitive expectations arises, triggering fear (MacDorman & Chattopadhyay, 2016). Although similar studies have further supported the uncanny valley theory and offered other reasonable explanations, significant controversies remain. These controversies mainly include three points: First, comfort as a dependent variable lacks a unified definition and effective measurement. Second, anthropomorphism as an independent variable is influenced by multiple factors, making it difficult to clearly define and systematically manipulate. Third, there is a lack of explicit mathematical models to fit the curve described by the uncanny valley theory (Wang, Lilienfeld, & Rochat, 2015). Understanding the essence of the uncanny valley theory and the boundary conditions for the emergence of uncanny feelings can effectively guide the design of social robots, preventing negative impacts on human-robot interaction caused by excessive anthropomorphism.

4.2 Media Equation Theory

Researchers have found that when people interact with inanimate objects, they unconsciously apply social norms from interpersonal interactions to these objects, even though they know the objects are inanimate. This theory is called the media equation (Reeves & Nass, 1996). Experimental evidence supports this theory: when experimenters used computers to provide information to participants, participants perceived computers with female voices as more caring and those with male voices as more knowledgeable (Nass, Moon, & Green, 1997). This aligns with everyday gender stereotypes—that women are generally considered to have better caring abilities while men have stronger agency (Huddy & Terkildsen, 1993). Not only with computers but also with robots, people unconsciously attribute social norms to them. An experiment conducted at a university showed that nearly 50% of participants greeted a robot receptionist in some manner before engaging in deep interaction, just as they would with a real person. This greeting predicted that subsequent interactions with the robot would be more social and polite (Min, Kiesler, & Forlizzi, 2010). The media equation theory also manifests in racial bias: for two robots with identical appearance and function, German participants rated a robot with a German name and manufactured in Germany as having better design and performance and as being more intimate than another robot described as having a Turkish name and manufactured in Turkey (Eyssel & Kuchenbrandt, 2012). These results demonstrate that in human-robot interaction, people do not merely treat robots as tools but also endow them with certain social attributes, further complicating the human-robot binary relationship.

4.3 Mind Perception Theory

Do robots have minds? Alan Turing, one of the earliest researchers to address this question, designed the famous Turing test: if an interrogator cannot distinguish between human and machine responses, the machine is considered to have human intelligence (Turing, 1950). Human minds are generally thought to consist of two components: agency (such as self-control, emotion recognition, planning, communication, thinking, moral character) and experience (such as joy, anger, worry, sadness, fear, surprise) (Gray, Gray, & Wegner, 2007). Mind perception refers to humans perceiving varying degrees of mind in robots during interaction. Gray et al. (2007) elucidated this theory, finding that robots were considered to have limited experiential capabilities but moderate agency capabilities, indicating that most people sense some form of mind in robots. Even though current technology cannot enable robots to have self-awareness like real humans, at least in the eyes of those interacting with them, they possess minds and thoughts to some extent.

5. Ethical Considerations and Research Prospects

The famous American science fiction writer Isaac Asimov (1950) proposed the Three Laws of Robotics in his work: First, a robot may not injure a human being or, through inaction, allow a human being to come to harm. Second, a robot must obey orders given it by human beings except where such orders would conflict with the First Law. Third, a robot must protect its own existence as long as such protection does not conflict with the First or Second Law. These laws originate from science fiction but have practical significance. Since robots are products of human intelligence, do they have rights like humans? A study on attitudes toward guide robots found that when children were with their parents, they were friendly toward robots. However, when unsupervised by adults, children might gather to abuse robots, such as blocking their paths, insulting and kicking them, and ignoring the robots' polite requests to stop (Brscić, Kidokoro, Suehiro, & Kanda, 2015). Designers should make reasonable decisions about robots' appearance and behavior to mitigate the risk of robot abuse. Furthermore, research findings in this field have reference value for understanding human behavior and reducing animal abuse risks. While humans demand absolute loyalty and obedience from robots, should they also consider their responsibility toward robots, such as respecting robots' work and properly recycling decommissioned robots? These moral issues in human-robot interaction deserve careful consideration by researchers and should be applied to the design and development of social robots.

Due to current technological limitations, although participants in existing human-robot interaction research paradigms are told they are interacting with robots, this is actually pseudo-interaction—most robots are operated by humans behind the scenes, similar to traditional Chinese puppet shows where puppeteers control the puppets' movements and voices from behind the curtain. This manipulation inevitably makes participants feel that robots' responses are stiff and unnatural. Consequently, humans may not have the same emotional experience when interacting with robots as they do with real people. Manufacturing robots requires overcoming numerous engineering challenges, and enabling robots to have human-like emotions, cognition, and behavior is an even more formidable challenge. Facing this challenge, we need to break the dominance of engineers and computer experts in artificial intelligence development and incorporate more disciplines, including psychology. Existing research themes in human-robot interaction require further deepening and expansion.

5.1 Trust Theme

As one of the important social signaling mechanisms, trust facilitates cooperation and establishes the foundation for sound social relationships by reducing social transaction costs. Research on trust in interpersonal interaction typically employs the trust game (TG) paradigm and uses brain imaging technology to explore the cognitive and neural mechanisms underlying trust behavior. Wang

Yiwen et al. (2015) used a repeated trust game (rTG) to explore the temporal dynamic characteristics of brain activity during trust interactions. The results showed that during the decision-making phase, distrust choices elicited more positive P2 components (150-250 ms) than trust choices. In the feedback phase, loss feedback elicited more negative FRN components (200-300 ms) than gain feedback, while gain feedback elicited shorter P300 latencies than loss feedback. Behavioral results also indicated that individuals chose trust significantly more often than chance level. Whether the same patterns of brain activity and behavioral results exist in human-robot trust interaction remains to be verified. Currently, few researchers have combined functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs) technologies in human-robot trust research to analyze the similarities and differences in cognitive and neural mechanisms of trust between human-human and human-robot interactions. Future human-robot trust research should fully utilize technologies such as fMRI, EEG, and ERPs that have been employed in interpersonal trust research, and design interaction scenarios that align with robots' practical applications to explore the cognitive and neural mechanisms of trust in human-robot interaction. By adjusting factors that may influence human-robot trust, such as the degree of anthropomorphism, we can enhance people' s trust levels in robots, enabling humans and robots to interact at moderate trust levels to achieve good interaction outcomes.

5.2 Empathy Theme

Empathy can regulate the formation and development of socially acceptable behavior, comprising two factors: emotional responses that share others' affective states, and cognitive factors that consider issues from others' perspectives (Decety & Jackson, 2004). For robots to exhibit socially acceptable behavior and integrate into human society, research on the emotional impact they produce is essential. Existing studies have used fMRI and EEG technologies to examine empathy differences when individuals interact with humans versus robots (Leite et al., 2013; Suzuki, Galli, Ikeda, Itakura, & Kitazaki, 2015). However, the empathy scenario priming paradigms used in these studies—such as robots commenting on chess players' moves or robotic hands being cut by scissors—have some distance from robots' actual applications in daily life. Future research should focus on robots' practical applications, exploring empathy issues in human-robot interaction within contexts such as hospital care, student education, and mall guidance, to discover designs more suitable for robots in these scenarios.

5.3 Group Identity Theme

In daily life, individuals often clearly distinguish interaction partners as in-group or out-group members based on certain labels. This distinction allows individuals to perceive whether interaction partners belong to the same social group, thereby influencing psychological processing and behavioral decisions during interaction. This theory is called social identity theory (Tajfel, 1978). Compared to out-group members, individuals are more willing to share outcomes and bear negative consequences with in-group members, show significantly higher emotional empathy for in-group members, exhibit more cooperation and reciprocity, reach agreements more easily, and pay more attention to fairness within in-group members. In interpersonal interaction, researchers have used the ultimatum game (UG) to examine how group identity affects responders' fairness concerns and their dynamic temporal processes, finding that group identity can influence individuals' early attention resource allocation and fairness processing (王益文等, 2014). In human-robot interaction, it remains to be explored whether robots will be categorized as out-group members by interacting individuals, and whether the influence of robots' group identity (in-group/out-group) on individuals' psychological processing and behavioral decisions is consistent with human-human interaction. Related research findings can help engineers manufacture robots that better fit into human groups.

Overall, research themes such as trust and empathy in human-robot interaction can be further explored in depth. Human-robot interaction research can refer to existing interpersonal interaction research themes and utilize interpersonal interaction paradigms such as trust games, chicken games, and ultimatum games for expansion and enrichment, to explore similarities and differences between human-robot and human-human interactions and reveal the psychological processing mechanisms and cognitive neural mechanisms underlying these interactions. On this basis, we can further examine the establishment, change, and influencing factors of the human-robot binary relationship, providing effective guidance for robots to better integrate into human society and for manufacturing more suitable robots for human society.

Most existing research focuses on the mechanical and software engineering levels of robots, with studies concentrating on human-robot interaction itself still relatively scarce. With the rapid development of artificial intelligence, robots' functions are becoming increasingly powerful. How to integrate highly intelligent robots into human society rather than merely making them showpieces will be an issue for future research to explore and resolve. In other words, the field of artificial intelligence needs the integration of psychology to form emerging interdisciplinary fields such as robot psychology or artificial intelligence psychology, helping artificial intelligence and its corresponding products develop in ways more acceptable to humans.

The emergence of artificial intelligence technologies such as robot nursing and drone delivery shows us the possibility of AI replacing humans in certain physical and mental labor tasks. Although AI can outperform humans in certain work domains (such as vehicle assembly and cashiering) through optimized algorithms and high-speed computing, it may still be unable to communicate with humans naturally and emotionally. Similarly, in fields requiring creative thinking and heuristic reasoning, AI may not yet have reached human levels. During the further expansion and optimization of AI technology, how humans and AI can coexist deserves careful consideration. The integration of psychology can help people better understand and solve the symbiotic problems between humans and AI. How existing AI technologies can be accepted and promoted by people, and how AI technologies still under development can be integrated into human customs and social norms, are precisely the directions that psychology needs to develop when integrated into the AI field. The intersection of psychology and artificial intelligence will enable humans to better meet the challenges of the AI era, while also allowing psychology to play its proper function in the process of AI changing the world. The development of science and technology should be human-centered, focusing on human needs. In the research and development of both artificial intelligence and robots, we should not merely focus on expanding their practical functions but should fully consider the importance of psychological factors, enabling these products to better integrate into human society and benefit humanity. Psychology will play a crucial role in AI cognitive structure design, AI learning to imitate human emotions and behavior, and AI-assisted therapy.

Acknowledgments

We especially thank Dr. Yin Jun from Ningbo University for his comments and assistance with this research.

References

[1] 邓卫斌, 于国龙. (2016). 社交机器人发展现状及关键技术研究. 科学技术与工程, 16(12), 163–170.

[2] 王益文, 张振, 张蔚, 黄亮, 郭丰波, 原胜. (2014). 群体身份调节最后通牒博弈的公平关注. 心理学报, 46(12), 1850–1859.

[3] 王益文, 张振, 原胜, 郭丰波, 何少颖, 敬一鸣. (2015). 重复信任博弈的决策过程与结果评价. 心理学报, 47(8), 1028-1038.

[4]Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. Games and Economic Behavior, 10(1), 122–142.

[5]Broadbent, E., Tamagawa, R., Kerse, N., & Knock, B. (2009). Retirement home staff and residents' preferences for healthcare robots. Paper presented at the Ro-Man 2009 - the IEEE International Symposium on Robot and Human Interactive Communicatio (pp. 645–650). Piscataway, NJ: IEEE.

[6]Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. Annual Review of Psychology, 68, 627-652.

[7]Brscić, D., Kidokoro, H., Suehiro, Y., & Kanda, T. (2015). Escaping from children' s abuse of social robots. Paper presented at the Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction -Hri (pp. 59–66). New York: ACM.

[8]Chang, C. W., Lee, J. H., Chao, P. Y., Wang, C. Y., & Chen, G. D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. Journal of Educational Technology & Society, 13(2), 13-24.

[9]de Graaf, M. M. A., Allouch, S. B., & Klamer, T. (2015). Sharing a life with Harvey: Exploring the acceptance of and relationship-building with a social robot. Computers in Human Behavior, 43, 1-14.
[10]Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. Behavioral and Cognitive Neuroscience Reviews, 3(2), 71-100.
[11]Diehl, J. J., Schmitt, L. M., Villano, M., & Crowell, C. R. (2012). The clinical use of robots for individuals with autism spectrum disorders: A critical review. Research in Autism Spectrum Disords, 6(1), 249-262.
[12]Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: anthropomorphism as a function of robot group membership. British Journal of Social Psychology, 51(4), 724-731.
[13]Frédéric Kaplan. (2001). Artificial attachment: will a robot ever pass ainsworth' s strange situation test. Paper presented at the Proceedings of second IEEE-RAS International Conference on Humanoid Robots (pp. 125-132). Piscataway, NJ: IEEE.
[14]Freedy, A., Devisser, E., Weltman, G., & Coeyman, N. (2008). Measurement of trust in human-robot collaboration. Paper presented at the International Symposium on Collaborative Technologies and Systems (pp. 106-114). Orlando, FL: IEEE.
[15]Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. Science, 315(5812), 619.
[16]Han, J., Park, I. W., & Park, M. (2015). Outreach education utilizing humanoid type agent robots. Paper presented at the Proceedings of the 3rd International Conference on Human-Agent Interaction (pp. 221-222). New York: ACM.
[17]Haring, K. S., Matsumoto, Y., Watanabe, K., Haring, K. S., Matsumoto, Y., & Watanabe, K. (2013). How do people perceive and trust a lifelike robot? Paper presented at the International Conference on Intelligent Automation and Robotics (vol 1, pp. 425 430). San Francisco, USA.
[18]Isaac Asimov. (1950). I, robot. United States: Gnome Press.
[19]Ishiguro, H., & Nishio, S. (2007). Building artificial humans to understand humans. Journal of Artificial Organs, 10(3), 133-142.
[20]Kim, Y., & Mutlu, B. (2014). How social distance shapes human-robot interaction. International Journal of Human-Computer Studies, 72(12), 783-795.
[21]Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., & Paiva, A. (2013). The influence of empathy in human-robot relations. International Journal of Human-Computer Studies, 71(3), 250-260.
[22]MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. Cognition, 146, 190-205.
[23]Mann, J. A., MacDonald, B. A., Kuo, I. H., Li, X., & Broadbent, E. (2015). People respond better to robots than computer tablets delivering healthcare instructions. Computers in Human Behavior, 43, 112-117.
[24]Min, K. L., Kiesler, S., & Forlizzi, J. (2010). Receptionist or information kiosk:how do people talk with a robot? Paper presented at the ACM Conference

on Computer Supported Cooperative Work (pp. 46-47). New York: ACM.
[25]Mori. (1970). The uncanny valley. Energy. 7, 33-35.
[26]Mubin, O., Stevens, C. J., Shahid, S., Mahmud, A. A., & Dong, J. (2013). A review of the applicability of robots in education. Technology for Education and Learning, 1, 209-215.
[27]Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. Journal of Applied Social Psychology, 27(10), 864-876.
[28]Park, R. E. (1924). The concept of social distance: as applied to the study of racial relations. Journal of Applied Sociology. 8, 339-344
[29]Phillips, E., Zhao, X., Ullman, D., & Malle, B. F. (2018). What is human-like?: Decomposing robots' human-like appearance using the anthropomorphic roBOT (ABOT) database. Paper presented at the Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (pp. 59-66). New York: ACM.
[30]Reeves, B., & Nass, C. I. (1996). The media equation: How people treat computers, television, and new media like real people and places. New York, US: Cambridge University Press.
[31]Robinson, H., Macdonald, B., Kerse, N., & Broadbent, E. (2013). The psychosocial effects of a companion robot: a randomized controlled trial. Journal of American Medical Directors Association, 14(9),
[32]Rosenthal-von der Pütten, A. M., Schulte, F. P., Eimler, S. C., Sobieraj, S., Hoffmann, L., & Maderwald, S., et al. (2014). Investigations on empathy towards humans and robots using fMRI. Computers in Human Behavior, 33, 201-212.
[33]Rosheim, M. (2006). Leonardo´s lost robots. Berlin: Springer.
[34]Sabelli, A. M., & Kanda, T. (2015). Robovie as a mascot: A qualitative study for long-term presence of robots in a shopping mall. International Journal of Social Robotics, 8(2), 211-221.
[35]Sandoval, E. B., Brandstetter, J., Obaid, M., & Bartneck, C. (2015). Reciprocity in human-robot interaction: A quantitative approach through the prisoner' s dilemma and the ultimatum game. International Journal of Social Robotics, 8(2), 303-317.
[36]Satake, S., Hayashi, K., Nakatani, K., & Kanda, T. (2015). Field trial of an information-providing robot in a shopping mall. Paper presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 1832-1839). Piscataway, NJ: IEEE.
[37]Suzuki, Y., Galli, L., Ikeda, A., Itakura, S., & Kitazaki, M. (2015). Measuring empathy for human and robot hand pain using electroencephalography. Scientific Reports, 5.
[38]Tajfel H. (1978). Differentiation between social groups: Studies in the social psychology of intergroup relations. London: Academic Press.
[39]Turing, A. M. (1950). Computing machinery and intelligence. Mind, 59(236), 433-460.
[40]Vallèsperis, N., Angulo, C., & Domènech, M. (2018). Children' s imaginaries of human-robot interaction in healthcare. International Journal of

Environmental Research & Public Health, 15(5), 970.
[41]Vincent, J., Taipale, S., Sapio, B., Lugano, G., & Fortunati, L. (2015). Social robots from a human perspective. Berlin: Springer.
[42]Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The uncanny valley: existence and explanations. Review of General Psychology, 19(4), 393-407.
[43]Wang, X., & Krumhuber, E. G. (2018). Mind perception of robots varies with their economic versus social function. Frontiers in Psychology, 9(1230).

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