# RoboPsychology: Developing Empathetic Robots

Prashant,22MCA20999
University Institute of Computing
Chandigarh University, Mohali (India)
prashantsharma5873@gmail.com

Amanpreet Kaur
University Institute of Computing
Chandigarh University, Mohali (India)

## 1. Abstract

Robotic psychology is a field of study that examines the interaction between humans and robots, and how humans perceive and respond to robots in social and psychological contexts. The field of robotic psychology encompasses a range of topics, including the design and development of robots thatare capable of engaging with humans on an emotional and social level, the impact of robots on human behavior and attitudes, and the ethical and societal implications of integrating robots into various aspects of daily life.

The study of robotic psychology draws on various disciplines, including psychology, cognitive science, neuroscience, and engineering. Researchers in this field seek to understand how humans form emotional connections with robots, the factors that influence trust and acceptance of robots, and the potential for robots to enhance human well-being and social interactions.

Robotic psychology has applications in a wide range of fields, including healthcare, education, entertainment, and industry. As robots become more sophisticated and pervasive, understanding how humans interact with and respond to them will become increasingly important for designing effective and beneficial robotic technologies.

## 2. Keywords

Socially assistive robots (SARs), Clinical robotics, Rob therapy, Robot in lifespan**.**

## 3. Introduction-

Artificial beings are now a common occurrence in o ur daily lives.
People are accustomed to using voice assistants like Siri, Google Assistant, and Amazon Alexa to suppo rt their daily activities, for instance.
Researchers have given some of these artificial entities a physical embodiment as a result of recent tech nological developments and a new paradigm in robotic sciences that emphasises the "human-oriented" values of engineering design, allowing the m to interact in a more typically human manner (e.g ., proxemics, kinesics, tactile and multisensory stimulation)
The development of personal service robots made them strong enough to be used in a variety of areas, including healthcare, where they have proven to be particularly useful. Yet, in order to effectively operate and fulfil the objectives of the healthcare interventions, personal service robots had to learn to interpret social information and communicate in a socially appropriate and sophisticated manner. They transformed into social robots in other terms.
According to both scientific and non-scientific criteria, social robots can communicate with and assist humans
More specifically, social robots are anticipated to, to varying degrees, satisfy a number of requirements.
Social robots must:

- Refrain from creating erroneous expectations based on their appearance, size, and material properties;
- Recognize, engage with, and use, whenever possible, all communication modalities that humans ordinarily employ; and
- Be cognizant of social rules and norms and act accordingly.

Robotic psychology is an emerging field of study that examines the social and psychological interaction between humans and robots. With the rapid advancement of robotics technology, the role of robots in our daily lives is becoming increasingly prevalent. As such, it is essential to understand the psychological mechanisms underlying our interactions with robots.

Robotic psychology draws on a range of disciplines, including psychology, cognitive science, neuroscience, and engineering. It explores topics such as human-robot interaction, emotional connections between humans and robots, the impact of robots on human behavior and attitudes, and the ethical and societal implications of integrating robots into various aspects of daily life.

The study of robotic psychology has significant implications for various fields, including healthcare, education, entertainment, and industry. For instance, in the healthcare field, robots are being developed to assist with patient care and provide emotional support. In education, robots are being used as tools for teaching and learning, while in the entertainment industry, robots are used to create interactive and immersive experiences for audiences. This research paper will provide an overview of the field of robotic psychology, including its history, theoretical foundations, and current research trends. It will examine the various factors that influence human-robot interaction, including personality traits, social norms, and cultural differences. Additionally, this paper will explore the potential benefits and challenges associated with integrating robots into different domains of human life. Overall, this paper aims to provide a comprehensive understanding of the field of robotic psychology and its implications for the future of human-robot interaction.

Interest in empathic robots is growing in academia and industry. Softbank's Pepper is designed to understand a human's mood and respond accordingly, which requires both emotion recognition and an expression engine. The long- awaited social robot Jibo was released in the market in 2018 with a range of social skills, including identifying family members and calling by names, telling jokes, and dancing. While interacting with such robots may certainly be entertaining, it is still early to say that state-of-the-art commercialized robots can empathize with humans.

We feel similar emotions as others, which is sometimes a result of understanding others' thoughts and feelings. Empathy involves "an effective response more appropriate to someone else's situation than to one's own". Empathy considers the other's affective state and situation, which leads to cooperation, prosocial behavior, altruism, and a positive relationship. It seems critical for robots to empathize with human partners, that is, recognize human emotional states, thoughts, and situations, and behave accordingly in order to live with human partners at home, to help with their mental or health-related problems, or to assist their daily activities.

Researchers of human–robot interaction (HRI) have recently started exploring different aspects of empathy. The current state of research is far from achieving full-fledged empathetic capability, but recent progress in social and developmental psychology, neuroscience, and virtual agent research have highlighted research directions for empathic social robots.

The purpose of this review is twofold: to understand the effects of robotic empathy on humans and to identify the components necessary to design and engineer empathy for robots. Apparently, these two understandings inform each other. Researchers may also identify gaps in research and gain insights into establishing aresearch agenda.

To this end, we systematically selected literature on empathy in interpersonal, human–agent, and human–robot interaction. We then provide a working conceptual model of empathy applicable to HRI, based on the literature on social, developmental, and clinical psychology and neuroscience. This model comprises empathy processes, outcomes, and modulator factors of empathy. We then review the literature on virtual humans and social robots to extend our model of empathy. While the conceptual model of empathy is certainly not a computational model for empathic robots, it may provide a blueprint for a cognitive–affective architecture for engineers.

A general design guideline for empathic robots will be provided to inform designers about the elements required to engineer empathic robotics. Our review identifies that, depending on the purpose, context, and tasks of a social robot, critical factors of empathy to implement may vary. We outlined three types of empathic robots as a function of the complexity of the empathic process.

## 4. Methods

We referred to the most recent meta-analysis on social robots by Duradoni et al. [**7**] when establishing a search strategy for a systematic analysis. We limited the search terms to include "empathy" in conjunction with critical keywords related to interpersonal interaction ("dyadic", "social", "interpersonal"), and human–agent interaction ("embodied conversational agent", "virtual humans", "avatars", "agents"), and human–robot interaction ("social robots", "HRI", "robots"). We defined our inclusion criteria to be literature that is: a paper published in a peer- reviewed journal or a conference proceeding, written in English, published until 2021, anempirical study.

We used databases of Google Scholar, Psych Articles, Psych Info, PubMed, Science Direct, and Sociological Abstracts. **Table 1** includes the number of articles considered in our systematic review. The initial screening of the abstract resultedin 1116 (interpersonal interaction), 128 (human– agent interaction), and 76 (human–robot interaction). We excluded 188 articles based on the following exclusion criteria: (1) Interpersonal literature: the definition of empathy is identical to the original study, the paper does not add substantial findings to the literature, the article applies only to a limited domain; (2) Human–agent and Human–Robot literature: the manuscript adopted a loose folk-definition of empathy, the research investigates whether the participants empathize with the system and not an empathicsystem, the paper has critical flaws. As a result, we selected 70 (interpersonal interaction), 10 (human–agent interaction), and 12 (human–robot interaction) articles for a review.

**Table 1.** The number of articles screened, assessed,and included for a review.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

## 5. Empathy in Interpersonal Interaction

The origin of empathy can be traced to the German term Einfühlung, which connotes the observer's projection to the physical object of beauty. Lippes later adapted this concept to understand other people. The English term empathy was coined by Titchener as atranslation of Einfühlung.

Empathy research has been conducted in the fields of social, developmental, and clinical psychology, and later neuroscience. Since the discovery of mirror neurons in monkeys, neuroscientists have identified underlying neurological evidence for empathy. Overlapping brain patterns were observed when an observer perceived the same emotions from a target, suggesting shared affective neural networks.

However, there is no consensus on the definition of empathy. The number of definitions is proportional to the number of researchers. Scholars agree that empathy consists of multiple subcomponents. A few critical elements of empathyare commonly identified across definitions. This review organizes prominent views on empathy in interpersonal research and suggests a comprehensive definition of HRI. Our definition has two functional roles: deciding which empathy literature to include or exclude for our review and establishing the cornerstones of the conceptual model of empathy.

The cognitive and affective aspects of empathy are probably the two most discussed topics in this field of study (see **Table 2**). Empathy definition are organized into three groups: definitions with emphasis on affective, cognitive, and both aspects of empathy. Only the original definitions are considered; that is, definitions that are mechanically combined are excluded. From this point on, all critical elements of empathy that merit their inclusion in the conceptual model are in italics.

**Table 2.** Definitions of empathy.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

Many researchers argue that empathy has two components: affective and cognitive. Affective empathy generally connotes the observer's visceral reaction to the target's affective state. Cognitive empathy involves taking the target's perspective and drawing inferences about their thoughts, feelings, and characteristics. Several researchers exclude or conditionally include cognitive aspects in the definition of empathy. For example, Zaki claimed that perspective taking, a cognitive process, is only regarded as a part of empathy when it involves experience sharing. Proponents of a narrower view also argue the difficulty in pinpointing the nature of the automaticity of empathy with the inclusion of cognitive components.

In our research, we adopted an inclusive approach and acknowledged both aspects of empathy. This is to avoid confining HRI research to only motor mimicry or emotion contagion and to include extensive empathic interaction based on cognitive elements such as context, past experience, and knowledge about the dyadic. In other words, to establish rapport and relationship with a human partner, a social robot requires both affective and cognitive aspects of empathy. As such, recent HRI research incrementally includes research on perspective taking, a form of cognitive empathy, of social robots.

As shown in **Table 2**, except Davis, most definitions stress the outcome of empathy. While most define empathic responses as similar or congruent, a few narrower definitions denote feelings identical to those of the target. The empathic outcome is certainly a result of the empathizer's or observer's internal empathy process. A clear division of process and outcome is critical for specifying the causal relationship between the two. This specificity is required to engineer empathy through the cognitive architectureof a social robot.

One important aspect of empathy in HRI is its purpose. Social robots are designed with a particularpurpose in mind, so the architecture of the empathic process should be designed to serve its goal. Given this, we define empathy for HRI as the robot's (observer) capability and process to recognize the human's (target) emotional state, thoughts, and situation and to produce affective or cognitive responses with the purpose of eliciting a positive perception of humans. The human perception of an empathic social robot ranges from liking, trust, and intention for long-term use, which we will address later.

A review of the current literature resulted in a working conceptual model of empathy (see **Figure)**

This model outlines the processes and critical components of empathy that are applicable to the design of social robots. It is an intermediate model to be evolved, with a review of empathy research on virtual agents and HRI. Construct elements and interaction processes are identified as comprehensive to include representative empathy scenarios involving social robots.

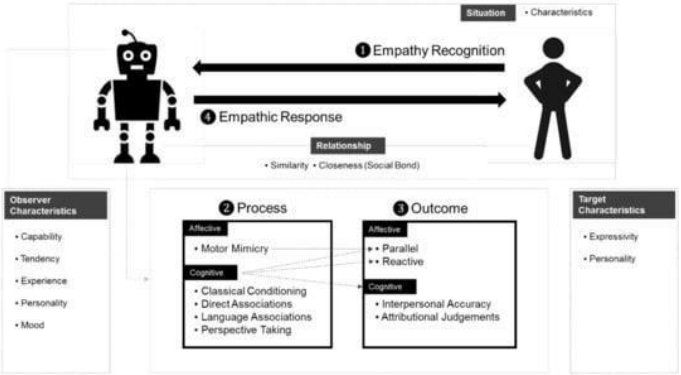


**Figure 1.** Conceptual model of empathy of HRI.

The assumption is that for humans to perceive a robot's empathy positively (e.g., increased liking and trust), the robot's empathy should be engineered similar to humans' processing of empathy. As robots and humans are essentially different, elements of human empathy that have little value in HRI were excluded. Therefore, we

emphasize that this is not an integrated model for understanding human interpersonal empathy, but rather a selective and organizing model to design an empathic social robot.

A typical empathic episode is initiated when the observer (robot) perceives empathic cues (expression or situation) from the target (human) through verbal or nonverbal. The observer then engages in an internal affective or cognitive process. This results in the observer's internal outcomes. The observer may decide to express the emotional state or behavior to the dyadic target. If done so, an empathic response is given to the target. We deliberately separated and as first suggested by Davis in his interpersonal empathy model. This includes scenarios in which a robot empathizes with a human but keeps the emotions to itself, deciding not to express them at the moment.

The characteristics of and the relationship between the observer and target, and the situation as to where, when, and what kind of empathic event occurred are the modulating factors that influence the processes. A detailed account of each element follows, with an explanation of the process and outcome preceding the empathy recognition and response.

### 6.1. Processes

Empathy processes are the underlying mechanisms that produce empathy outcomes. We integrated and identified the most prominent empathy theories and organized them into affective and cognitive processes. Each process has differential neural straits. The two mechanisms merit different routes for empathy in an emotion–cognitive computational module because they may lead to different empathic outcomes.

**Motor mimicry**. This refers to the observer's automatic and unconscious imitation of the target. Mimicry was first described by Lipps and organized by Hoffman into a two-step process:

the observer imitates the target's empathic expressions (e.g., facial expression, voice, and posture); this imitation results in afferent feedback that produces a parallel effect congruent with the target's feedback, as depicted in **Figure 1** For example, a robot may imitate a human's facial expression, who looks cheerful, and changes its emotional state accordingly. This mechanism is also referred to as primitive emotional contagion or the chameleon effect. Mimicry is important in building rapport and makes the observer more persuasive; however, under certain situations, this may produce a diminishing effect.

**Classical conditioning**. occurs when a neutral stimulus (NS) is repeatedly paired with an unconditioned stimulus (US), leading to an unconditioned response (UR). Once conditioned, the sole conditioned stimulus (CS) is sufficient for the observer to exhibit a conditioned response (CR). Similarly, empathy can occur when

the observer pairs the empathic cues (NS) of the target with his or her emotional cues (US) and the associated affective state (UR). Such cues are not limited to the target's facial expressions, but also include the situation and context in which empathic interactions occur.

For example, a family may have several members, with each member following their different schedules. Nevertheless, when they gather together in the evening or on weekends, they have a joyful time, full of positive emotions. A social robot may learn this connection between family members gathering together (NS), the positive facial expressions of family members (US), and the corresponding positive emotions for a robot (UR). Once conditioned, a social robot may expect and prepare services congruent with family gatherings (taking photos and dancing).

However, conditioning is probably the least studied empathy process in the HRI.

**Direct associations.** When the observer perceives the target's empathic cues, the observer feels the emotions attached to it if they match the observer's past experience. This is the general version of classical conditioning. Social robots may have episodic memories with associated emotions and use them to "feel" the current situation. For example, a robot may visually recognize two people hugging and then draw from its past experience involving a hug that includes warm, nurturing, and calm emotions.

Language associations. Sometimes, empathy is the result of a language-based cognitive network that triggers an observer's emotional state.

Language-mediated association does not require direct observation and is considered a more

advanced cognitive process. Eisenberg et al. explained a similar process, dubbed an elaborated cognitive network.

This process typically involves a conversation or dialogue with a social robot. A target human may tell the robot that he or she went to a party the previous night. The word "party" may trigger the robot's language network and its past emotions in a party. Technologically, language-mediated associates require the social robot to have automatic speech recognition (ASR) and natural language processing (NLP), and a semantic map of words associated with emotions.

**Perspective taking**. Perspective taking or role taking is considered the most advanced cognitive process among empathy processes. This involves the observer's effortful process of imagining the target's perspective and suppressing the observer's perspective. This advanced cognitive process alongside language-mediated association is what many researchers call cognitive empathy.

A robot should project an imaginary situation and state of the observer to mimic perspective taking, which humans do with considerable effort. For example, one fine morning, a robot may greet a human target who has a tired look on the face and may ponder why this person looks tired. It may consider several reasons; for example, it may consider that this being exam week, the human target may have underslept, which may cause it to state: "I hope you are not jeopardizing your health. You may be stressed about the exam, but sleeping well is also necessary to study well." This is an empathic concern (reactive outcome). All of this requires a virtual construction of the target's situation and the emotional states associated with it from the robot's side.

The difference between the first three (motor mimicry, classical conditioning, direct associations) and the latter two (language associations, perspective taking) categories is whether the observer directly perceives the empathic cues. The latter two can be invoked without directly perceiving the target's empathic cues.

### 6.2. Outcome

The empathy process of an observer yields affective or cognitive outcomes. Affective outcomes are further divided into parallel and reactive outcomes. Parallel outcome indicates the matching of the observer's emotion to the target's and has been the focus of early empathy research. The matching of emotion denotes an affective outcome congruent with the target, as suggested by many researchers. Motor mimicry or classical conditioning may lead to a parallel outcome.

Empathy outcomes sometimes go beyond similar or congruent reactions that do not reproduce the observer's state. Reactive outcome involves the observer's emotion, which is different from the targets. Feelings falling under this category include sympathy, empathic concern, empathic anger, and personal distress.

Outcomes may be primarily cognitive, such as interpersonal accuracy—that is, an estimation of the target's thoughts, feelings, and characteristics.

Insights are also regarded as the product of the cognitive empathy process, typically by perspective taking. This ability depends on the empathic capability of the observer. In counseling psychology, the result of perspective taking does not necessarily imply emotional ties with the client.

Another important aspect of interpersonal accuracy applicable to HRI is its anticipatory nature. A social robot may project many imaginary scenarios constructed from the target's empathic cues, with each scenario anticipating the target's behavior. A social robot may weigh in each scenario considering factors such as context and past history of the observer and may suggest helpful services to the human observer (e.g., turning off lights and playing classical music).

Much more technological advancement for a social robot is required to make attributional judgments, which refer to identifying the causes behind the target's thoughts, feelings, and characteristics. Human empathizers generally attribute causes to the observer's situation rather than attributions. Human empathizers tend to identify dispositional attributions as a cause of the target's success and situational attributions for the target's failure.

### 6.3. Observer and Target Characteristics

Many factors modulate the empathy process. First, the characteristics of the observer, target, and relationship between the two play an important role.

Empathy requires an observer's empathic capability and the ability to perform the empathic process. This requires all elements of **Figure 1** to recognize empathic cues, process empathy, and express empathic responses to the target. This is a stable characteristic of human observers. We included such characteristics for HRI because of its evaluative value. Similar to the Turing test, an empathy test for a robot is required to calibrate a social robot for empathy capability. This test identifies the strengths and weaknesses of many aspects of empathy and reveals whether the robot is appropriate for a specific social task in a certain setting. Designers and engineers may view this as the goal to be attained. To the best of our knowledge, there is currently no measurement to evaluate a robot's empathic capability.

However, just because the observer is capable does not necessarily mean that he or she is likely to empathize with an empathic event. Empathic tendency involves an individual's predisposition to engage in empathic interactions. This is a stable characteristic for humans, and self-reported measures have been developed to evaluate it. For a social robot, we view empathic tendency as a more fluid variable to be manipulated. In the movie Interstellar, the robot TARS's humor level could be changed by its human counterpart. Given that the need for empathy depends on the situation, the willingness to empathize or not has its own merit. This is especially important for social robots that are expected to interact with individuals in different situations (i.e., domain-independent), which we see as the final step of an empathic social robot (i.e., Type III empathic robot).

The accuracy of empathic response can only be ensured when targets' expressivity enables their thoughts and emotions to be perceived. In other words, a social robot cannot empathize if the human user does not express emotional cues For designers of HRI, this means that empathy scenarios have to be carefully designed so that human users are led to express such cues without giving the impression that one is forced to do so (e.g., directly asking how one feels).

Past experiences of the observer are relevant because many cognitive empathic processes relate to past memory. The HRI research continues to maintain and operate a memory-like structure for empathy, even though their domains and scenarios are limited.

The mood and personality of observers (robots) are also important modulating factors. Given that the result of empathy affects one's emotion, a social robot may have an emotion module embodying the empathy process. There is a close relationship between emotions and personality. Emotions are temporary, and personality remains stable over a long period of time. A robot's personality has long been researched yet, to our knowledge, no interaction with empathy exists.

Finally, there is a clear female superiority in empathic capability due to raising and nurturing children, and a recent HRI study has revealed a few interesting observations with differential gender effects in interaction with a social robot.

### 6.4. Relationship

Characteristics such as similarity and social bonds are a joint function of both the target and observer. The stronger the observer–target similarity, the stronger the likelihood and the intensity of the observer's empathic response Specifically, Cialdini et al. demonstrated that empathy stemming from similarity is related to a sense of self–other overlap, an emotional signal of oneness. Additionally, observers tend to empathize more with targets with similar personalities and values. Familiarity is also suggested to modulate the strength of an empathic response.

Given their similarity and familiarity, humans tend to be more empathic toward individuals with whom they have social bonds, including friends and family members, than strangers. The type of relationship affects the type of prosocial action taken by the observer. Social robots may bear similar empathy to certain events but behave differently with people according to relationships.

### 6.5. Situation

All empathic interactions occur within a specific situational context or behavior. In the HRI, this indicates the type of robot's tasks and goals involving the empathic processes as well as where and when such interactions occur. The

observer's contextual appraisal has been argued in interpersonal research and physiopsychologically analyzed, suggesting several modulatory factors.

Not all empathic events were equally treated. Davis emphasizes the strength of the situation, which influences the power of empathic responses in interpersonal interactions. A helpless target suffering from a traumatic event tends to produce a powerful empathic outcome. A social robot may tag priorities on empathic events and modulate empathic responses accordingly.

Empathy is also moderated by the target's behavioral characteristics. Based on neurological evidence from a functional magnetic resonance imaging study, Singer et al. found that men empathize more with people with fair social behavior. Lamm et al. revealed that observers had a smaller empathic response to pain-afflicted targets when cure was justified.

### 3.6. Empathic Recognition

If a social robot is viewed as an information processor, empathic recognition is the input stage, and the empathic response is the output stage. This stage involves the characteristics of empathic cues and the type of modality used to recognize such cues. Humans utilize a full range of interpersonal modalities, including verbal and nonverbal communication channels. The verbal channels primarily involve speech or dialogue, which directly—but not exclusively— connect to language-based cognitive networks in the empathic process. Nonverbal channels involve gait, posture, gestures, movement, distance, facial expressions, gaze, and physical factors (e.g., touch), as well as nonSemitic voice features (e.g., speech speed, pitch, variation). Many lead exclusively to motor mimicry.

Through such modalities, the target releases empathic cues. We empathize more with primary emotions (e.g., happiness and sadness) than with secondary emotions (e.g., jealousy). Strong negative cues may elicit stronger observer responses. This is also introduced as the intensity and saliency of the observed emotion. For example, the intensity of the target's pain-induced facial expressions modulates the observer's empathic response. This is due to the increasing or decreasing attention to an empathy-eliciting stimulus. The effect is further moderated by target characteristics (e.g., a helpless person).

### 6.7. Empathic Response

Empathic interaction involves a constant parallel interaction between recognition and response. Empathic responses are directed towardthe target, and the factors are nearly identical to recognition, including the modality and extent to which empathy is expressed.

However, higher-level constructs should be addressed, such as the purpose of expressing empathic responses. One purpose may be to build rapport and improve social relationships with human partners. On the other hand, a social robotmay not be interested in building a long-term relationship but aim at increasing liking and trust through helping and caring behavior. As a social robot operates for a specific purpose, task goals should be outlined to select the modalities and expressions to adopt during interactions.

Empathic interaction is initiated (**Figure 1**) when the observer recognizes the target's empathic cue or behavior in a specific contextual situation. The observer, the target's characteristics, and the relationship modulate empathic processes and outcomes. The nature of a social robot's empathic response depends on its relationship with the human target and the situational context of interaction. Our framework clearly distinguished between the outcome and empathic response so that while a robot may "feel in" and bear an empathic outcome depending on modulating factors, it may calibrate the strength of an empathic response or even decide not to express one.

## 4. Empathy in Human–Agent and Human–Robot Interaction

Research on empathy with virtual humans or embodied conversational agents (ECAs) has been conducted with robots because it is less difficult to implement and manipulate empathy. Research on both virtual humans and robots, dubbed "advanced intelligent systems" when combined, focuses on either of the following two perspectives: (1) humans' empathic response to the advanced intelligent system or (2) the effect of a robot's empathic behavior on humans. The former does not

necessarily involve robots that have empathic capabilities but focuses on how humans empathize with robots that have human-like characteristics. As our study is primarily interested in the latter, we only considered original studies that designed and implemented empathy in an advanced intelligent system.

**Table 3** summarizes the seminar studies on empathic virtual agents. Given that all advanced intelligent systems are designed with predetermined goals, we organized the literature working backward, from the virtual agent's goal or the study's purpose. We then identified key elements identified in our empathy framework (**Figure 1**), such as observer and target characteristics, the relationship between the two, and the situation involving empathic interaction.

**Table 3.** Studies on empathic virtual agents.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

Empathic virtual agents have been studied in the context of playing games, healthcare interventions, job interviews, email assistance, social dialogue, or even story narration.

A typical empirical study investigated the effects on participants' perceptions when interacting with or observing an empathic virtual agent compared with a non-empathic one. Overall, empathic agents were perceived positively in liking and trust and felt more human-like , caring , attractive, respected , and enjoyable . Some caveats were revealed, such as participants' negative perceptions of agents' incongruent empathic responses and participants being stressed to an empathic agent. While most studies were based on one-time interaction, a few studies identified the participants' intention to use empathic virtual agents longer. The research community certainly has established a grounding that an empathic virtual agent, when implemented to provide an appropriate response congruent with the situation, elicits the positive perception of users for long-term interaction.

A few recent studies have taken the next step to investigate the effects of empathy modulating variables, such as the observer (virtual agent)'s mood, personality, and relationship (similarity, familiarity, liking, social bond) . However, such manipulation was applied only to virtual agents interacting with themselves.

The empathic tendency, which our model defined as an individual's predisposition for an empathic response, was implemented as an empathic threshold for a virtual agent to respond to the target's empathic event . Although the studies demonstrated that the higher the weighted means of modulating factors, the higher the level of the virtual agent's empathy, the research did not examine the interaction effect between the factors being modulated. That is, we do not have evidence of how liking and familiarity interact. The Boukricha study assigned weighted values to each variable but used the sum of such values, assuming an additive effect. A carefully designed study investigating an empathic agent with modulation compared with an empathic agent with no modulation with human participants is required. We now take a closer look in **Table 4** at the empathic processes of each study.

**Table 4.** Empathic recognition and responses of virtual agent studies.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

Empathic cues were recognized from the participants' affective states measured by facial recognition or physiological measures. For example, the empathy model presented by Boukricha et al. assumed an emotional event when a fast and salient change occurred in the emotional state of the virtual agent. Almost all studies considered the user's situation elicited from, for

example, their multiple dialogue choices when considering the participant's affective state.

The most sophisticated empathic processes, integrating both affective and cognitive empathic processes, were demonstrated by Rodrigues et al. The candidate emotion of the target was elicited from the facial expression (motor mimicry) as well as the projected emotion elicited from the appraisal of the situation (perspective taking). The two emotions were compared to produce a potential empathic emotion. However, the study was limited to a virtual agent emphasizing another virtual agent; therefore, follow-up studies may be conducted with human interaction.

In a strict sense, virtual agent studies only involved motor mimicry and perspective taking, but not classic conditioning, direct associations, and language associations, as outlined in **Figure 1**. That is, none of the studies have demonstrated conditioning scenarios (i.e., classical conditioning). No empathy research had the virtual agent store its experience and reference the agent's past experience to elicit the emotion to them (i.e., direct associations). No study involved language-based cognitive networks engineered to demonstrate a virtual agent hearing a word and then eliciting certain emotions by triggering language-mediated associations (i.e., language associations). In short, most studies with empathic virtual agents have focused on motor mimicry and perspective taking, so a major push is required to extend our understanding of empathy in advanced intelligence systems.

**Table 5** presents the seminal studies on the effects of social robots evoking empathy. Empathic social robots have been researched only in a limited domain, such as playing games or quizzes, healthcare, or a human's one-way utterance to or a conversation with a robot. Many studies involving playing games utilized Philip's iCat robot because of its ability to express facial expressions, but most recent studies utilized Softbank's Pepper due to its multimodal capability.

**Table 5.** Studies on empathic robots.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

Many studies on empathic social robots have game playing situations such as chess. Chess is selected probably because both the participant and the robot can take turns so that the conditions can be easily tracked. Robots can also express empathic emotions based on the situation of the game. Most importantly, the participant may clearly recognize the robot expressing empathy during turns.

Compared with virtual humans, research on social robots is at the stage of infancy, and participants' perceptions of empathic robots are limited to perceived trust, human-likeness, friendliness, social presence, and engagement. However, the most recent HRI research applied an advanced computational model (e.g., deep learning) to model empathy, which merits more discussion in **Section 5**. We now take a closer look at the empathic processes of each study in **Table 6**.

**Table 6.** Empathic recognition and responses in social robot studies.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

Interestingly, while most studies on both virtual agents and empathic robots, involving motor mimicry, had an intelligent system to respond with the same modal (i.e., facial mimicking), Hegel's study involves cross-modal motor mimicry. That is, when participants read an emotional story to a robot, it responded with a congruent emotional expression on its face. The idea of cross-modal mimicry is grounded by Chovil. More sophisticated research will face the question of how to combine empathic cues from different modalities and combine them into a singular representative value.

We defined this cross-modality empathic feature as one of the key features of the Type III Social Robot.

Overall, research on empathic virtual agents and social robots seems to elicit a human's positive response when a robot's empathic response is expressed and the outcome is congruent with the situation. Perceived measurements that may gain people's interest in long-term interactions, including liking, trust, attractiveness, engagement, enjoyment, believability, and human-likeness, had a positive effect with an empathic virtual agent or a robot. Surprisingly, except for Cramer et al., none of the studies measured perceived empathy directly. We suggest referencing established perceived empathy measures, such as the relationship inventory measures by Barrett-Lennard. The relationship inventory measures possessed adequate internal reliability, and the scales fit the kind of interaction with a social robot. A few scales can be slightly modified to the HRI, such as "the robot tries to understand me from my point of view", a scale measuring cognitive empathy, or "the robot does not understand the way I feel".

## 7. Synthesis and Discussion

Developing an empathic robot, like all other robots, requires a concrete function definition that serves the purpose of the robot. Based on such functions, designers may design interaction scenarios, and engineers may develop the software and hardware architecture of the robot.

One consideration is the purpose of the robot. Is the robot to assist passengers at airports? Is it a companion robot built for long-term interactions? Is it a counseling robot that provides health-related advice? This purpose defines the level and type of empathic capabilities. The demanded capability requires a number of empathic characteristics.

We grouped the three types of empathic robots that organized such characteristics (see **Figure 2**). Generally, the more applicability in terms of domains and tasks, more sophisticated empathic robots (i.e., Type III robots) are required. Currently, most empathic research on social robots is slowly moving to Type II robots, and we are starting to see Type I empathic robots in the industry. As HRI research precedes commercialization, we expect to see research on Type III robots in the near future.

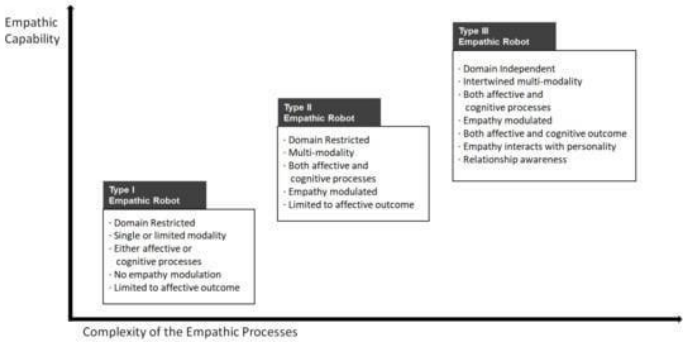


**Figure 2.** Empathic capability as a function of the complexity of empathic processes.

We do not suggest a full-fledged empathic robot (Type III) for all situations. Certain tasks may not require sophisticated empathic capabilities. Depending on the tasks, users may perceive empathic responses negatively. The exact areas in which users require empath is an important research question to be illuminated.

As depicted in **Figure 2**, the following factors need to be considered when designing an empathic robot. As research on empathic social robots is at the stage of infancy, certain factors may lack empirical evidence.

### 7.1. Domain-Dependent vs. Domain-Independent

All social robots in the literature are heavily domain-dependent. For example, a social robot may be exclusively designed for educational purposes. A robot cannot function outside the designed application. That is, currently there is no empathic robot that can switch tasks and service multiple domains. For example, we are yet to see an empathic robot that serves coffee at work and then interacts with family members after work. As reviewed, the empathic event may influence the likelihood of empathizing and the power of empathic responses. A domain-independent robot should seamlessly manipulate its empathic tendency (i.e., a robot's predisposition to engage in an empathic interaction) according to the situation at hand.

One reason why domain-independent empathic robots (i.e., Type III empathic robots) are difficult to produce is that, with the current AI technology involving voice recognition and natural processing, all use cases should be defined and trained for a robot to recognize. Designers should carefully

describe all possible interaction scenarios as well as the kind of empathy process to engineer. For example, should we engineer empathy for a robot that serves coffee? If so, to what extent? Some customers may not necessarily expect empathy.

### 7.2. Emotion Modeling

Our definition of empathy for social robots involves the robot's (observer) capability and process to recognize the human's (target) emotional state, thoughts, and situation and to produce affective or cognitive responses. Although there are exclusive cognitive responses involving interpersonal accuracy and attributional judgment, almost all resulted in affective responses, either parallel or congruent. Only Prendinger's study involved an exclusive cognitive response.

As empathy involves emotions, designers should define the kind of emotion that may occur in a given scenario for both the observer and the target. If emotion recognition and expression require sophistication, an emotion model is required. Existing research is not clear regarding which model is more effective or may gain a more positive perception of the social robot by the participant.

### 7.3. Single vs. Multimodality

Most studies have focused on single modal interaction with empathic robots, but we are now seeing an increasing number of multimodal robots. Multimodal recognition is powerful because the robot recognizes redundant cues for a better understanding of the user's affective state and thoughts. Owing to situational reasons, users tend to choose one modality over another. If the robot does not support such a channel, the opportunity to recognize its empathic cue is lost. The most advanced Type III robot may combine multimodal channels to maximize empathic interactions. This is what humans do as well.

### 7.4. Empathy Modulation

As we have reviewed in **Figure 1**, given the same empathic cues, the empathy outcomes differ by modulating variables including the situation (strength and characteristics) and relationship (similarity, familiarity, liking, and social bond). Designers should possess a good understanding of tasks and contexts. Are there multiple users involved? Should the empathic robot respond differently to different users with various relationships? When empathic cues differ in strength, how should empathy change?

### 7.5. Affective and Cognitive Outcomes

Although nearly all social robot research is limited to either an affective outcome (parallel response to the target) or a cognitive outcome, many virtual agents have virtual agents for producing both outcomes. We want to emphasize that the outcome is not necessarily expressed to users. For example, an empathic robot may feel a certain way that is represented by its emotion model but decide not to exercise restraint in its feelings because of certain reasons; humans do this all the time. The Type III empathy robot should engineer this control module that moderates its emotions and empathic expressions.

### 7.6. Interaction with Personality

Currently, there is no research on the effect of differential personality on the empathic capability of a social robot. However, emotions are intertwined with personality. As we have seen in the case of AI assistants, more customers require customization of their assistants (e.g., voice personas), so there is a need for different personalities to modulate the empathic process.

### 7.7. Anthropomorphic versus Biomorphic

According to Bartneck and Forlizzi , social robots can be classified as anthropomorphic (i.e., mimicking a human) or biomorphic (i.e., mimicking a lifelike object). Interestingly, the two most used robotic platforms for an empathic study are biomorphic (i.e., cat-like) iCat and anthropomorphic Pepper. We have not seen any studies on how different forms affect empathy. Future studies may address this.

### 7.8. For a Hybrid Computational Model

Yalcin and DiPaola recently reviewed computational models in artificial agents. They argued that the data-driven approach to model empathy is in its infancy. This is consistent with our findings that few HRI studies on deep learning-based computational models were published only recently in 2020 and 2021, respectively.

Meanwhile, conceptual models based on empathy theories would provide a useful framework because they outline mechanisms involving empathy and include solid explanatory power. Eventually, we

expect a hybrid computational model where data are gathered, analyzed, and predicted (i.e., data-driven) based on the sub-components envisioned by the theory-driven model.

In this article, we reviewed all prominent empathy theories and established a conceptual framework that illuminates critical components to consider when designing an empathic robot, including the empathy process, outcome, and the observer and target characteristics. This model is complemented by empirical research involving empathic virtual agents and social robots. We identified many gaps in the current literature on empathic social robots and overviewed essential factors to be included in a computational model. We also suggest that critical factors such as domain dependency, multi-modality, and empathy modulation be considered when designing, engineering, and researching empathic robots.

**REFERENCES-**


1. Pollak, A.; Paliga, M.; Pulopulos, M.M.; Kozusznik, B.; Kozusznik, M.W. Stress in Manual and Autonomous Modes of Collaboration with a Cobot. Comput. Hum. Behav. **2020**, 112, 106469. [**Google Scholar**] [**CrossRef**]
2. Wainer, J.; Feil-seifer, D.J.; Shell, D.A.; Mataric, M.J. The Role of Physical Embodiment in Human-Robot Interaction. In Proceedings of the ROMAN 2006—The 15th IEEE International Symposium on Robot and Human Interactive Communication, Hatfield, UK, 6–8 September 2006; pp. 117–122. [**Google Scholar**]
3. Haring, K.S.; Satterfield, K.M.; Tossell, C.C.; de Visser, E.J.; Lyons, J.R.; Mancuso, V.F.; Finomore, V.S.; Funke, G.J. Robot Authority in Human-Robot Teaming: Effects of Human-Likeness and Physical Embodiment on Compliance. Front. Psychol. **2021**, 12, 625713. [**Google Scholar**] [**CrossRef**] [**PubMed**]
4. Holland, J.; Kingston, L.; McCarthy, C.; Armstrong, E.; O'Dwyer, P.; Merz, F.; McConnell, M. Service Robots in the Healthcare Sector. Robotics **2021**, 10, 47. [**Google Scholar**] [**CrossRef**]
5. Van Wynsberghe, A. Service Robots, Care Ethics, and Design. Ethics Inf. Technol. **2016**, 18, 311–321. [**Google Scholar**] [**CrossRef**][**Green Version**]
6. Wan, S.; Gu, Z.; Ni, Q. Cognitive Computing and Wireless Communications on the Edge for Healthcare Service Robots. Comput. Commun. **2020**, 149, 99–106. [**Google Scholar**] [**CrossRef**]
7. Sarrica, M.; Brondi, S.; Fortunati, L. How Many Facets Does a "Social Robot" Have? A Review of Scientific and Popular Definitions Online. Inf. Technol. People **2019**, 33, 1–21. [**Google Scholar**] [**CrossRef**]
8. Bartneck, C.; Forlizzi, J. A Design-Centred Framework for Social Human-Robot Interaction. In Proceedings of the RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759), Kurashiki, Japan, 22 September 2004; pp. 591–594. [**Google Scholar**]
9. Chen, S.-C.; Moyle, W.; Jones, C.; Petsky, H. A Social Robot Intervention on Depression, Loneliness, and Quality of Life for Taiwanese Older Adults in Long-Term Care. Int. Psychogeriatr. **2020**, 32, 981–991. [**Google Scholar**] [**CrossRef**]
10. Costescu, C.A.; Vanderborght, B.; David, D.O. The Effects of Robot-Enhanced Psychotherapy: A Meta-Analysis. Rev. Gen. Psychol. **2014**, 18, 127–136. [**Google Scholar**] [**CrossRef**][**Green Version**]
11. Libin, A.V.; Libin, E.V. Person-Robot Interactions from the Robopsychologists' Point of View: The Robotic Psychology and Robotherapy Approach. Proc. IEEE **2004**, 92, 1789–1803. [**Google Scholar**] [**CrossRef**