Human Experience And Robotic Experience: A Reciprocal Exchange Of Perspectives

Sara Incao*, Carlo Mazzola*, Francesco Rea, Alessandra Sciutti, Massimo Marassi

Abstract

The basis of a humane approach to others is the authentic comprehension of another subject. As humans, we can achieve this understanding well, and the reason lies in how we experience the world around us and other subjects in it. The development of robots capable of socially interacting and helping humans is progressing, even though they are still far from reaching an autonomous comprehension of others' intentions, emotions, and feelings. In this sense, the humane approach may be addressed in robotics through the concept of experience. In a reciprocal exchange of perspectives, the core elements and the structure of human experience are investigated, together with how the idea of experience has been implemented in robots. The embodied Self and the relationship with other subjects form the pivot of any human experience and are suggested to be the basis for the emergence of a more profound idea of experience in robots. Conversely, the possibility to develop a robot with a primitive sense of Self raises questions about the nature of human experience and the impact such technologies have on it.

The experience as a core element of a novel interdisciplinary exchange

Cognitive robotics inspired by human beings

In robotics, a specific branch aims to endow robots with models inspired by the most advanced cognitive agent known: the human being. The objective of Cognitive Robotics is to develop and unify systems of perception, action, memory and learning that might provide robots with higher autonomy and flexibility when interacting with the environment (Vernon, 2014; Sandini et al., 2019). One of the merits of this branch is to unite the practical implementation of cognitive and neuroscience research outcomes on a robotics platform equipped with a physical artificial body. This approach stands to become even more intriguing, considering humans more as subjects of experience than as cognitive systems. Indeed, to have an experience or to experience something is part of our daily life as human beings. We have experiences about the world when we taste chocolate, touch the frozen snow, or look at the colorful feathers of a tropical parrot. However, we can also experience our painful muscles after a long swim or butterflies in the stomach before presenting at an important conference. These experiences, exteroceptive, proprioceptive, and interoceptive respectively, are inherently tied to the body and allow us to interact in the environment and with others.

Robotic experience

When dealing with artificial robotic systems, the term ‘experience’ acquires a specific connotation. Robotic platforms are endowed with multiple kinds of sensors through which the robot retrieves information from the environment. Moreover, robots can also continuously collect information about their inner state: for example, the position of their joints and the velocity of their movements. All these data can be stored in memory and processed by algorithms targeted to make the robot situated in the environment, to optimally plan the robot's actions or train the robot on a specific task.

For specific applications, where the context and task are well defined and fixed – such as, for example, traditional automotive production – robots pre-programmed to reproduce the same operations accurately may represent an effective solution. Conversely, a robot meant to navigate everyday life in society needs to be endowed with the ability to learn and adapt to the ever-changing environment and the unpredictability of human beings (Vernon et al., 2016). For this reason, the application of machine learning algorithms to robots represents a massive field of research with significant results, although highly task-dependent. For example, a particular technique used to train robots is Reinforcement Learning. Here, positive and negative rewards received in continuous interactions with the world reinforce the learning process. In this way, the robot can «learn from its experiences», i.e., find and learn about patterns in the environment and any related information regarding the world around it.

Experience from a first-person perspective

The sense of the word 'experience', as targeted by robotics, is in line with the operational aspect of human experience. Indeed, the methods we employ to face our daily interactions with the world involve our body in a combination of perception and action (Varela et al., 2016). Since childhood, our bodily experience has been the background of every learning process (Martin & Schwartz, 2005; Falck-Ytter et al., 2006). The experience of the world is structured by methods and patterns that we memorize and repeatedly reuse, adapting them to any occurrence. However, even though this operational aspect of experience is implicit in the everyday life of human beings, the meaning of the term 'experience' in humans invokes something even more articulated as a concept. Indeed, any experience refers to a specific subject possessing a personal or subjective character that concerns the first-person perspective of the individual. Every experience is therefore typical for the person to whom it belongs. It originates as related to the intrinsic motivations of that subject, colored by their affective states, intertwined with other experiences, structured in their moral beliefs.

Experience and meaning

In this sense, human experience is what allows us to interpret and give meaning to the world from a specific and unique perspective, so that every object (which might be something in the environment or another person) is perceived with reference to ourselves: our motivations, expectations, emotions, personality, memory, body. Therefore, whereas robots can already detect, store and learn information from the environment to solve a task, this ability is still far from the complexity of the human experience of the world. Furthermore, cognitive skills are necessary to devise meaning from the information processing of autonomous agents. The information coming from the environment through the senses gains meaning from motivations connected to the organism 's maintenance of balance. Damasio explains this process through the concept of homeostasis (Man & Damasio, 2019). The motivation behind human actions and behaviors cannot always be traced back solely to the need for survival. In fact, the scale of values we use to act and behave implies affective, instinctual, and moral incentives. Reference to the Self and to the maintenance of a sort of equilibrium is therefore crucial for human beings. Through this constant contact with their individual point of view, human beings can overcome the limitations of pure data received from the outside, by interpreting them in light of motivations, values, affections, memories and moral beliefs. Moreover, it is only through this constant allusion to an individual coherence that humans can establish a hierarchy in the flow of experiences, which otherwise would be a set of equivalent information.

Promotion of the dialogue

Albeit linked to the specificity of human experience (lacking in robots), dialogue around what makes an experience possible might be highly relevant in both the robotic and philosophical fields. Investigating the structure of human experience might inspire novel approaches to the improvement of robotic autonomy and social abilities (Gaggioli et al., 2021). On the other hand, discussion about the advances and potential developments of artificial embodied systems might promote a deeper investigation of the nature of human experience and the impact such technologies have on it.

A philosophical investigation of experience: the priority of interpretation

Definition of experience

As a preliminary and broad definition, *experience* outlines the way humans internalize the reality they live in. The word stems from the Latin *ex-periri*, which means to go through a trial, to attempt something. Therefore, it describes a form of knowledge concerning a consistent and repetitive situation that produces the conception and the understanding of a State of Things, the direct relation of the subjects to the otherness they are affected by, the ability to remember and organize impressions, the possibility to broaden knowledge through a verification process.

Experience as repetition

In all its forms, experience implies a series of events that appear to the subject who experiences them. The temporal sequence in which they appear includes the presence of elements that are recognized as analogous. Hence, the primary feature that describes experience is repetition. Then, backward reflection about the flow of different experiences allows the uniqueness of each event to be identified.

In fact, it might appear paradoxical that such uniqueness becomes evident only through the constant repetition of a phenomenon. However, it is only in this way that its nature may become known and its conditions of possibility identifiable. Moreover, only through this process is it possible to anticipate a novel event and verify its evidence. That being the case, the repetition of events might produce a twofold effect. The first is the prediction of a phenomenon not yet experienced but analogous to previous ones. The second is the identification of a uniformity of phenomena that can be formulated as a general law. More precisely, on the one hand prediction relies on repetition, while on the other, uniformity – once defined as a general law – exceeds iteration and is irrespective of temporality.

Theory and observed data in science

Reflection on the connotation of experience in science might deepen the relationship between a single experienced event and the constant general law that expresses its meaning. The parallel in the scientific method would be the structure of the experiment on the binomial of observed data and scientific theory. As Kant pointed out, explicitly referring to the Newtonian physics and Galilean astronomy of his time, no hypothesis could be deduced from observation (Kant, 1787). Therefore, the experiment is not a step towards creating a theory; it does not genetically precede the idea, but is rather the expression of fundamental questions that humans pose nature through theoretical hypotheses. Kant saw that the history of science confuted the emergence of a theory as a derivation from repeated observations. As he advocated, knowledge does not emerge from facts and phenomena but from *a priori* forms that build and order them. Moreover, as Popper expounded, the necessity and universality of theories do not arise from experiments. Experiments are instead suggested and interpreted by theories (Popper, 1962).

The priority of lifeworld

In light of this comparison with science, it is essential to explore a distinctive aspect of human experience by moving from an opposition. On the one hand, scientific experience is characterized by the pursuit of objectivity achieved through a method. On the other hand, human experience is marked by historicity, by the novelty manifested in the timing of events, and by the unifying nature of consciousness that collects them as a teleologically organized continuity of memory, intentionality and forethought.

Experience, as intended in the scientific method, needs to be confirmed and validated by repetition: *ubi non reperitur instantia contradictoria* (Bacon, 1623; Stuart Mill, 1882). However, this connotation of experience does not imply the idea of historicity: it is far from being an existential experience. Before any process of abstraction and generalization, a subject is situated in the lifeworld. The existential experience does not proceed by progressive abstractions, since any phenomenon appears against a background of subjective interpretation. Therefore, the process of experience does not move from the observation of a fact toward its generalization. Instead, the perception of a single event emerges from motivations, personal experiences and bodily dispositions that interpret its meaning and move toward the consequent teleologically oriented action execution.

The meaning of denial

There is another point that demonstrates the difference between experience as a process of generalization/abstraction and experience as an interpretative process of events: this is the different meaning of denial in science and life. In scientific experiments, if the hypothesis gets disconfirmed, the error can be amended by formulating a new theory. In human experience, denial is called blame, error or evil i.e., something which is not entirely identifiable with the universality of a concept since it refers deeply to the inner state of the subject. Thus, human experience takes place in single events interpreted in light of the subject's history: previous delusions and faults, current motivations, and prospects for the future. It is far from being a repetitive process in which every case is reduced to the universality of its generalization.

Bodily experience and interpretation

The cornerstone of every human experience is the body. The body is indeed the only means through which a subject is open to a relationship with the environment (Merleau-Ponty, 1945). However, it is essential not to consider it a mere sensory apparatus that allows data coming from the senses to enter the conscience. The perceived object is never neutral for the subject, it is never experienced as mere information. Instead, it is always perceived relative to subjectivity and directly interpreted by the body. In this sense, the body is not inanimate and lifeless. It is rather living and lived. Lived because it collects, like tree bark, the imprints that the flow of time leaves on it. Living because it continuously receives sensory stimuli from the external environment that strikes it in a specific – its own specific – way. Perception is thus the bodily relationship between the subject and the object, the primary form of human experience where the external world appears to the subject as already interpreted, meaningful, relative to the Self.

Self-awareness and bodily experience

Experience of the body

Having a body and *being a body*. These two expressions disclose the twofold nature of the body: that of being simultaneously experienced as an object and as a subject. As a matter of fact, the body differs from all other objects present in the world. It is always part of us, and we cannot part with it or live without it. The expression *having a body* highlights the body's role as the origin of our contact with the world. The body we possess can sense the world and is the means by which we interact with the sensory world.

For this reason, the idea of *having* a body alludes to the experience of an entity we can feel, sense, touch, think about, conceptualize and take as the object of our thoughts. The uniqueness of the object body is that it is experienced in the first person and therefore becomes part of a subjective perspective. Coupled with the feeling of *having a body* is the experience of *being that body*, i.e., the awareness of it as the condition of possibility for any imaginable experience. It is myself as the body I am that lives and interacts, making any experience possible.

Body Schema

Artificial bodies are supposed to integrate with the algorithmic *brain* of robots in performing tasks that humans usually carry out. A specific field of robotics research is focused on addressing this integration (Vernon et al., 2016), supporting the view that cognition is «deeply dependent upon the characteristics of the physical body of an agent»

(Wilson & Foglia, 2011). Doing things like walking across a room avoiding obstacles is a simple action for an adult human being, but it may be slightly more difficult for a child and even more complex for a robot. A lot depends upon the extent to which our body schema has grown accustomed to the world outside. Many authors have used the concept of body schema since the beginning of the last century, but it has sometimes suffered from the lack of any conclusive definition. However, as Gallagher (1986) points out, the origin of ambiguities lies in the doubts regarding the conscious or unconscious nature of the body schema besides the body image. For this contribution, it is helpful to follow Gallagher's conclusion. He explains that the body schema, in contrast to the body image, is not a conscious percept through which we think about our body as the object of our thoughts. The body schema is, rather, the non-conscious constant responsible for our body's operative performance in the environment. When we are about to cross a road or receive a ball during a volleyball match, the postures we assume and the movements we perform are caused by the organization of the body schema. It contains information such as the space occupied by the body or the arm's length, which determines the estimations necessary to grasp or avoid objects and many other details about the functional and operative performance of the body within the world. The concept of body schema has introduced a modality with which human beings build a sort of model or map of the world outside that is strictly linked to the experience of their own body. Indeed, this connection results in a direct and mainly unconscious contact with the environment.

The body and the Self interpret the world

The interactions with the world that form specific human experiences, are characteristic to each individual not only because the shape of human bodies differs, but because the experience of the world differs considerably between distinct human beings. The human body, and consequently the body schema, are inherently tied to the subjective perspective of everyone's specific experience. Human beings interpret the world based on a personal horizon of experience, which is the sum of motivations, beliefs, intentions, and actions towards reality. Every new experience from the outside takes shape internally, always in reference to the flow of our previous experiences. This background of subjective participation in the world refers to the concept of Self.

Artificial Self

The progressive resemblance with humans that artificial embodied systems are pioneering is not only at the level of physical appearance but also in terms of behavior during interaction with the environment. From this perspective, the individual unity that we experience as human beings, the Self, has been recognized as a core element, if robots are to gain autonomy and adaptivity during interactions. As a result, several studies on cognitive robotics are looking at the possibility of building robots with a sense of Self. This element is intended to be included in a cognitive architecture, a computational model of the structure of a human mind connecting different modules such as memory, perception, decision-making, action and learning, so as to confer functional autonomy to the robot. In particular, the question is whether the Self should be a fixed module, a sort of metacognitive unity in which information is conveyed and analyzed, or if the Self should rather be seen as an emergent feature that progressively develops and rises through the learning of new data.

The Self as a module or an emergent element

Since the early 2000s, robotics research has attempted to improve the cognitive and relational capabilities of robots. Self-awareness has been identified as the key element to develop such capabilities (Lee, 2020). Many studies have attempted to define and build a model of the Self to be added to the cognitive architecture.

In 2005, Kawamura et al., (2005) attempted to develop a robot with a sense of Self by using a cognitive architecture with three memory systems that are repeatedly accessed to monitor the robot's representations of its internal states and the progress of the task required. In their experiment, the robot was able to suspend the task it was carrying out to pay attention to a new stimulus identified as more urgent.

Also, Novianto & Williams (2009) have defined Self-awareness in robots to be characterized by the ability to redirect attention towards internal states. Here the focus is on the attentional process that is «the allocation of perceptual resources to analyze a subset of the surrounding world to the detriment of others» (Ferreira & Dias, 2014). Following Novianto & Williams (2009) the attentional process should be directed towards internal states to develop self-awareness in a robot.

In contrast, in Birlo & Tapus (2011) the Self is configured as an independent unit that monitors different memory modules, namely *buffers,* and decides which one requires attention. Additionally, in Chatila et al. (2018) the cognitive architecture is designed in modules. Each one of these modules is a decision-making system that implements one way of producing actions. In addition to this system, there is a meta-element called *meta-controller* responsible for analyzing every decision-making process and selecting the one in charge of controlling the robot at any given time.

All these studies show that self-awareness has been considered a crucial element for robots to interact with humans in as natural a way as possible.

Furthermore, other studies, inspired by developmental psychology, have attempted to let the Self emerge with unsupervised learning mechanisms. Starting from data acquired from the environment, the robot could generate emergent representations. The concept of *emergence* emphasizes the capability of a system to show properties or abilities that are not intrinsic to its components. Hoffmann et al. (2018), for example, proposed an approach to study how the humanoid robot, iCub, could form a representation of its body surface by receiving several stimuli on its artificial skin with capacitive sensors. Here, the topographic representation is not pre-scripted but emerges from the robot's interaction with the surrounding environment. Furthermore, Lanillos & Cheng (2018) developed a perceptual computational model for multisensory robots to derive their body configuration through sensory information.

These implementations of Self modules are more focused on developing a concept of self-awareness as direct knowledge of the internal states. The approaches inspired by developmental psychology are, instead, closer to the embodied cognitive processes humans experience in the first stages of their lives, when a primary and unconscious sense of Self emerges in the womb. Both of these approaches, addressing the topic from different perspectives, may be instrumental in providing efficient and operative solutions to improve the adaptivity and autonomy of robots.

Perception and intersubjectivity

Perceiving others through our body

A noticeable difference lies between perceiving the environment and perceiving other subjects in it. In both cases, our senses receive numerous impressions, but the body participates in perception differently. A deep connection exists between others' bodies and one's own so that the body is doubly involved. Others' movements, actions and even bodily reactions are received with sight, hearing and touch but are experienced by passing through one's own body schema, reflecting and interpreting the other's bodily experience. This is also suggested by the function of the mirror neuron system, a cerebral organization of neurons that activates both when one is performing an action or when one is merely seeing others performing the same action (Rizzolatti et al., 1996; Rizzolatti & Craighero, 2004).

Twofold maturation of Self and perception of others

After all, the strict ties with others' bodies are evident since the prenatal stage, when the fetus grows, is fed and takes oxygen within the maternal womb. This strong relationship with the mother remains similarly intense through the first months of life. From birth, the newborn understands that other persons are 'like me'. Only from the ninth month do they start considering others as other intentional agents, an understanding they will gradually develop over the years (Tomasello, 1999). Finally, in early childhood, this process reaches a critical step. The child starts distinguishing their mental states from those of others, avoiding the mere egocentric attribution of their mental states (Moll & Meltzoff, 2011; Tomasello, 1999). During this process of distinction, the Self emerges. The Self is not indeed a static, already formed structure since birth. It instead comes to consciousness and develops in a continuous comparison with others, that are gradually recognized as other different selves having other diverse experiences. Such a twofold maturation – the emergence of the Self and recognition of other selves – persists throughout life and shows that, whilst the Self develops in relation to others, others are perceived and their experiences interpreted based on one's own Self. Therefore, the perception of others emerges from an analogical process based on oneself, one's own experiences, and one's own body.

Primitive skills of understanding others in humans and robots

Since social interaction is one of the ten main challenges in robotics (Yang et al., 2018), cognitive and developmental robotics aims to provide robots with the social ability to understand other agents' states (Sciutti & Sandini, 2017). In this direction, robots have been endowed with the primitive social skills infants acquire during the first months or even weeks of life, such as the predisposition to focus on biologically moving objects rather than non-biological ones (Simion et al., 2008; Vignolo et al., 2017). Also, the ability to discriminate an averted from a direct gaze, and the preference for the latter (Farroni et al., 2002), as well as to detect the direction of the gaze, have been implemented in artificial embodied systems (for example, see Schillingmann & Nagai, 2015; Palinko et al., 2016).

Development of social skills in humans and robots

All these abilities are at the basis of higher cognitive and social skills along the developmental process. Gaze direction detection enables joint attention mechanisms, which underlie perspective-taking and higher forms of understanding others' intentions (Tomasello et al., 2005; Moll & Meltzoff, 2011). Perspective-taking is indeed the ability to grasp the visuospatial perspective of another person that is different from the egocentric one, and gather what is seen by them, and how (Salatas & Flavell, 1976; Flavell et al., 1981). Such an ability seems to occur both as a mentalizing process and as an implicit and automatic mechanism (Surtees & Apperly, 2012). In robotics, such skill has been developed on artificial platforms to improve collaboration with other agents (Trafton et al., 2005; Zhao et al., 2016) and to enable robots to better map the environment in relation to their human partners (Fischer & Demiris, 2016). Furthermore, in infancy, the process matures up to the development of a Theory of Mind (Baron-Cohen, 1995, 2000) that allows children to understand the beliefs, intentions, and desires of other individuals. This same cognitive skill was found to be fundamental for robots involved in social interactions at the dawn of modern robotics (Scassellati, 2002) and was later implemented in social robots as an adaptive skill (Bianco & Ognibene, 2019).

The body is a means to interpret others

The child's developmental process shows the fundamental role of sociality on the emergence of the child’s Self. The Self grows in a continuous comparison with others so that the social context, which will see the child acquiring language skills and symbolic thought, starts earlier, shaping and building the infant’s Self. Indeed, the process of distinction from the mother and the progressive formation of others' mind consciousness manifest the body's role in developing self-other distinction. As the neural evidence of mirror neurons also suggests (Rizzolatti & Craighero, 2004), the function of the body, and more precisely the body schema, also persists into adulthood. When perceiving other persons, they are interpreted through one's own body as a lens to experience the environment and other agents. Hence, one's own body is the means to interpret what is perceived of others. Personal body schema, previous experiences, affective states, instincts, motivations lie in the body and are essential to comprehending others, which are experienced and interpreted through one's own embodied Self. Therefore, since the same ability to understand others would be crucial also for robots collaborating with humans (Sciutti & Sandini, 2017), robots should be provided with a sense of embodied Self (Prescott & Camilleri, 2019), serving as a pivot to understand others.

The emergence of self-other distinction skills in robotics

The tight connection between one's own body and perception of others has been explored in cognitive robotics with interesting results. A fundamental ability allowing the understanding of others is the self-other distinction. Zhang & Nagai (2018) proposed a computational model for learning the internal bodily states of the robot that brings together proprioceptive feedback from the robot's joints and visual feedback from its cameras. In this way, the robot can predict its own proprioceptive feedback starting from observing itself, an ability that would simulate the imaginary body schema. Moreover, the skill of self-other distinction can emerge from the robot learning it can visually distinguish its own actions from those of others. Following the same inspiration, Lanillos et al. (2020) developed a learning algorithm that allows the robot to recognize itself in a mirror by recognizing some simple actions generated by itself, with the consequent ability to distinguish its own movements from the ones generated by others.

Learning to understand others' intentions in robotics

The ability to understand the intentions of a partner and anticipate their movements is essential for collaboration. The cognitive architecture, Hammer (Demiris & Khadhouri, 2006), was inspired by the mirror neuron system to provide robots with the ability to understand others' actions. In that architecture, the robot uses its motor control system, both for executing actions and for perceiving the actions performed by other agents.

Moreover, Vinanzi et al. (2019) proposed an artificial cognitive architecture for intention prediction. The idea, achieved with an unsupervised learning model, is based on detecting others' skeletons during their action and inferring the subsequent postures up to the end of the action. Copete et al. (2016) introduced a computational model for action prediction based on the co-development of action prediction and action production. The sensorimotor information gathered during action production, which is missing in others' action observation (tactile and motor feedback), were reconstructed based on the information learnt in action production to facilitate the other's action prediction.

Cognitive robotics research is therefore implementing social abilities in robots, which are inspired by the developmental process of children. As evidenced in the paragraphs above, robots have already been endowed with some of these abilities and can learn a number of social interaction skills. Perhaps, what may be missing is a stronger focus on the Self as a mean to understanding others, an element which is at the basis of all human experience.

Conclusive remarks

The basis of a humane approach to others is the authentic comprehension of another subject. Every behavior, inspired by this motivation, proceeds from a prior understanding of what others are feeling, their emotions and intentions. A fundamental requirement for a humane robot is, therefore, to understand others (Sandini & Sciutti, 2018). As humans, we can achieve that understanding only through our own embodied Self: we comprehend others starting from ourselves – an ability which infants develop over the years. Technological advances have already provided robots with many social skills. Attempts to provide robots with a primitive sense of Self and distinction from others have been conducted by leveraging machine learning techniques and starting from bodily interactions with the world around them.

To continue in this direction, we believe that the idea of experience outlined in this chapter needs also to be considered for robots, something that can be obtained only through a multidisciplinary dialogue. Cognitive robotics might indeed receive inspiration from the human way of experiencing the world and others. The concept of an embodied Self as a pivot of experience is cardinal and evidences the constitutional relationship between what is experienced and the subject of experience. As artificial cognitive systems, robots should therefore be developed in light of the concepts of Self and experience. The Self is not reduced to the ability of being aware of one's own internal states. It can reach a conscious level with reflection, but it is also the subject's flow of experiences. In this sense, it can be viewed as the reference and interpreting medium for any possible experience appearing in the flow. To develop a subjective experience, then, robotics needs to undergo a mindset shift: from the integration of information to the interpretation of such information on the basis of the embodied Self. Interpretation means that every piece of information is read as related to the embodied Self: its body schema, motivations, affective states, previous perceptions and personality. Such an Artificial Self would provide robots with a cognitive-experiential structure inspired by humans and would make them more autonomous in learning, and more able to understand other agents and establish a commonality with them.

On the other hand, there are also prolific perspectives on this topic linked to the philosophical thought behind research into the specific nature of human cognition and experience. The building of robots that can socially interact with humans, learn from the environment, and develop a primitive sense of Self raises questions about human nature and its difference from artificial systems. It is also crucial to explore the impact that novel artificial systems may have on humans. Indeed, social robots may mediate and influence the way humans interact with the world around them, as every technological device does (Ihde, 1990).

From this perspective, any research into human experience and the development of robots with a sense of Self and capable of experiencing others, should cooperate and proceed in parallel with the aim of a human-centric technology.

References

Bacon, F. (1623). De Augmentis scientiarum.

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. MIT Press / Bradford Books.

Baron-Cohen, S. (2000). Theory of mind and autism: A review. International Review of Research in Mental Retardation, 23, 169–184. https://doi.org/10.1016/s0074-7750(00)80010-5

Bianco, F., & Ognibene, D. (2019). Functional advantages of an adaptive Theory of Mind for robotics: a review of current architectures. 2019 11th Computer Science and Electronic *Engineering (CEEC)*, 139–143. https://doi.org/10.1109/CEEC47804.2019.8974334.

Birlo, M., & Tapus, A. (2011). The Crucial Role of Robot Self-Awareness in HRI. *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI),* 115–116. https://doi.org/10.1145/1957656.1957688.

Chatila, R., Renaudo, E., Andries, M., Chavez-Garcia, R. O., Luce-Vayrac, P., Gottstein, R., Alami, R., Clodic, A., Devin, S., Girard, B., & Khamassi, M. (2018). Toward self-aware robots. *Frontiers Robotics AI*, *5*(AUG). https://doi.org/10.3389/frobt.2018.00088.

Copete, J. L., Nagai, Y., & Asada, M. (2016). Motor development facilitates the prediction of others' actions through sensorimotor predictive learning. *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob),* 223–229. https://doi.org/10.1109/DEVLRN.2016.7846823.

Demiris, Y., & Khadhouri, B. (2006). Hierarchical attentive multiple models for execution and recognition of actions. *Robotics and Autonomous Systems*, *54*(5), 361–369. https://doi.org/10.1016/j.robot.2006.02.003.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878–879. https://doi.org/10.1038/nn1729.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *PNAS*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999.

Ferreira, J. F., & Dias, J. (2014). Attentional mechanisms for socially interactive robots - A survey. *IEEE Transactions on Autonomous Mental Development*, *6*(2), 110–125. https://doi.org/10.1109/TAMD.2014.2303072.

Fischer, T., & Demiris, Y. (2016). Markerless perspective taking for humanoid robots in unconstrained environments. *Proceedings - IEEE International Conference on Robotics and Automation*, *2016-June*, 3309–3316. https://doi.org/10.1109/ICRA.2016.7487504.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1-Level 2 distinction. *Developmental Psychology*, *17*(1), 99–103. https://doi.org/10.1037/0012-1649.17.1.99.

Gaggioli, A., Chirico, A., di Lernia, D., Maggioni, M. A., Malighetti, C., Manzi, F., Marchetti, A., Massaro, D., Rea, F., Rossignoli, D., Sandini, G., Villani, D., Wiederhold, B. K., Riva, G., & Sciutti, A. (2021). Machines like Us and People like You: Toward Human-Robot Shared Experience. *Cyberpsychology, Behavior, and Social Networking*, *24*(5), 357–361. https://doi.org/10.1089/cyber.2021.29216.aga

Gallagher, S. (1986). Body Image and Body Schema: A Conceptual Clarification. *The Journal of Mind and Behavior*, *7*(4), 541–554.

Hoffmann, M., Straka, Z., Farkas, I., Vavrecka, M., & Metta, G. (2018). Robotic homunculus: Learning of artificial skin representation in a humanoid robot motivated by primary somatosensory cortex. *IEEE Transactions on Cognitive and Developmental Systems*, *10*(2), 163–176. https://doi.org/10.1109/TCDS.2017.2649225.

Ihde, D. (1990). *Technology and the Lifeworld. From Garden to Earth*. Indiana University Press.

Kant, I. (1787). *Kritik der reinen Vernunft* (II).

Kawamura, K., Dodd, W., Ratanaswasd, P., & Gutierrez, R. A. (2005). Development of a robot with a sense of self. *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation, CIRA*, 211–217. https://doi.org/10.1109/cira.2005.1554279.

Lanillos, P., & Cheng, G. (2018). Adaptive robot body learning and estimation through predictive coding. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4083–4090. https://doi.org/10.1109/IROS.2018.8593684.

Lanillos, P., Pages, J., & Cheng, G. (2020). Robot self/other distinction: active inference meets neural networks learning in a mirror. *ECAI 2020: Proceedings of the 24th European Conference on Artificial Intelligence*, 2410–2416. https://doi.org/10.3233/FAIA200372.

Lee, M. H. (2020). *How to grow a robot: Developing human-friendly, social AI*. MIT Press.

Man, K., & Damasio, A. (2019). Homeostasis and soft robotics in the design of feeling machines. *Nature Machine Intelligence*, *1*(10), 446–452. https://doi.org/10.1038/s42256-019-0103-7.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science*, *29*(4), 587–625. https://doi.org/10.1207/s15516709cog0000_15

Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Librairie Gallimard.

Moll, H., & Meltzoff, A. N. (2011). Joint attention as the fundamental perspective in understanding perspectives. In *Joint attention: New developments in psychology, philosophy of mind*.

Novianto, R., & Williams, M. A. (2009). The role of attention in robot self-awareness. *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 1047–1053. https://doi.org/10.1109/ROMAN.2009.5326155.

Palinko, O., Rea, F., Sandini, G., & Sciutti, A. (2016). A Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration. *IEEE International Conference on Intelligent Robots and Systems*, *2016-Novem*, 5048–5054. https://doi.org/10.1109/IROS.2016.7759741.

Popper, K. (1962). *Conjectures and refutations: the growth of scientific knowledge*. Routledge.

Prescott, T. J., & Camilleri, D. (2019). The Synthetic Psychology of the Self. In *Cognitive Architectures* (pp. 85–104). Springer.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. In *Annual Review of Neuroscience* (Vol. 27, pp. 169–192). https://doi.org/10.1146/annurev.neuro.27.070203.144230.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*(2), 131–141. https://doi.org/10.1016/0926-6410(95)00038-0.

Salatas, H., & Flavell, J. H. (1976). Development of Two Components of Knowledge. *Child Development*, *47*(1), 103–109. https://doi.org/10.2307/1128288.

Sandini, G., & Sciutti, A. (2018). Humane Robots — from Robots with a Humanoid Body to Robots with an Anthropomorphic Mind. *ACM Transactions on Human-Robot Interaction*, *7*(1), 5–8. https://doi.org/10.1145/3208954.

Sandini, G., Sciutti, A., & Vernon, D. (2019). Cognitive Robotics (in press). In M. H. Ang, O. Khatib, & B. Siciliano (Eds.), *Encyclopedia of Robotics*. Springer.
Scassellati, B. (2002). Theory of Mind for a Humanoid Robot. *Autonomous Robots*, *12*, 13–24. https://doi.org/10.1023/A:1013298507114.
Schillingmann, L., & Nagai, Y. (2015). Yet another gaze detector: An embodied calibration free system for the iCub robot. *IEEE-RAS International Conference on Humanoid Robots*, *2015-December*, 8–13. https://doi.org/10.1109/HUMANOIDS.2015.7363515.
Sciutti, A., & Sandini, G. (2017). Interacting With Robots to Investigate the Bases of Social Interaction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(12), 2295–2304. https://doi.org/10.1109/TNSRE.2017.2753879.
Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *PNAS*, *105*(2), 809–813. https://doi.org/10.1073/pnas.0707021105
Stuart Mill, J. (1882). *A System of Logic, Ratiocinative and Inductive 1843* (VIII). Harper & Brothers.
Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and Automatic Perspective Taking in Children and Adults. *Child Development*, *83*(2), 452–460. https://doi.org/10.1111/j.1467-8624.2011.01730.x
Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129.
Trafton, J. G., Cassimatis, N. L., Bugajska, M. D., Brock, D. P., Mintz, F. E., & Schultz, A. C. (2005). Enabling Effective Human–Robot Interaction Using Perspective-Taking in Robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *35*(4), 460–470. https://doi.org/10.1109/TSMCA.2005.850592.
Varela, F. J., Thompson, E., & Rosch, E. (2016). *The embodied mind: Cognitive science and human experience*. MIT press.
Vernon, D. (2014). *Artificial cognitive systems: a primer*. MIT Press.
Vernon, D., von Hofsten, C., & Fadiga, L. (2016). Desiderata for developmental cognitive architectures. *Biologically Inspired Cognitive Architectures*, *18*, 116–127. https://doi.org/10.1016/j.bica.2016.10.004.
Vignolo, A., Noceti, N., Rea, F., Sciutti, A., Odone, F., & Sandini, G. (2017). Detecting biological motion for human-robot interaction: A link between perception and action. *Frontiers Robotics AI*, *4*(JUN). https://doi.org/10.3389/frobt.2017.00014.
Vinanzi, S., Goerick, C., & Cangelosi, A. (2019). Mindreading for Robots: Predicting Intentions via Dynamical Clustering of Human Postures. *2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 272–277. https://doi.org/10.1109/DEVLRN.2019.8850698.
Wilson, R., & Foglia, L. (2011). Embodied cognition. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Summer 2021 Edition)*. https://plato.stanford.edu/archives/sum2021/entries/embodied-cognition.
Yang, G.-Z., Bellingham, J., Dupont, P. E., Fischer, P., Floridi, L., Full, R., Jacobstein, N., Kumar, V., McNutt, M., Merrifield, R., Nelson, B. J., Scassellati, B., Taddeo,

M., Taylor, R., Veloso, M., Wang, Z. L., & Wood, R. (2018). The grand challenges of Science Robotics. *Sci. Robot*, *3*, 31. https://doi.org/10.1126/scirobotics.aar7650.

Zhang, Y., & Nagai, Y. (2018). Proprioceptive Feedback Plays a Key Role in Self-Other Differentiation. *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob),* 133–138. https://doi.org/10.1109/DEVLRN.2018.8761042.

Zhao, X., Cusimano, C., & Malle, B. F. (2016). Do people spontaneously take a robot's visual perspective? *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 335–342. https://doi.org/10.1109/HRI.2016.7451770.