Hildelith Leyser and Kate Cooper***

Motion, Mirror Neurons and the Human–Robot Relationship

Abstract

The rapid advancement of robotic technologies and their integration into human environments necessitates a deeper understanding of the human–robot relationship. This article explores the dynamic interactions between humans and robots, particularly focusing on the trust mechanisms involved when humans interact with non-human agents. The attribution of human-like qualities to robots involves a complex interplay of cognitive processes, including anthropomorphism and motion perception. Drawing upon theories in cognitive science and recent empirical research, we examine how motion and perceived intentions play a critical role in forming human trust towards robots.
By exploring these relational dynamics, we aim to shed light on the potential of robots as active participants in human social settings, offering fresh perspectives on the roots of intersubjectivity and thus providing a framework for discussion of the ethical implications of trust in robots.

1. Introduction

The curious human tendency to ascribe something like human intention to non-human actors has long been recognised, and it is also well known that we do this not only with animals but also with inanimate objects. Famously, in an experiment conducted at

* Department of Neurology and Neurosurgery, McGill University
** Department of History, Royal Holloway, University of London

Smith College in the 1940s, Fritz Heider and Marianne Simmel recorded the pronounced tendency of study participants to construe not only intentionality but even an emotionally complex narrative when offered an intentionally simplified stimulus, an animated film showing two-dimensional geometrical shapes—two triangles and a circle—moving in white space.[1]

In the twenty-first century, the tendency to impute intention to non-human actors has ever wider ethical implications. As we move into the age of what Sherry Turkle has called the 'relational artifact'—the "computational object explicitly designed to engage a user in a relationship"[2]—non-human artefacts have begun to take on tasks previously fulfilled by human colleagues and carers, becoming ever more thoroughly integrated into human society. Human beings have begun to rely on these tools for cooperation as well as for help, safety and pastoral care. Establishing a smooth path for human–robot interaction (HRI) is now big business, with out-sized potential for both a positive and a negative impact. So it is important to understand what characteristics elicit a trusting response and—which is at least as important—what, if anything, can be done to ensure that the ability to elicit trust is accompanied by genuine trustworthiness.

Within this frame, the problem of trust is not straightforward. Turkle has referred to a 'crisis of authenticity', referencing the computer scientist Joseph Weizenbaum's disappointment with how student users reacted to Eliza, a natural language processing programme he created. Eliza used string matching and substitution to reply to statements with questions or re-statements of what the user had said, creating an effect not unlike that of a therapist trained in the non-directive method advocated by Carl Rogers.[3]

The source of Weizenbaum's disappointment was not his programme's failure, but rather its success. The programme had a positive emotional effect on those who used it, and this struck him as wrong. "Weizenbaum came to see students' relationships with Eliza as immoral", Turkle explains,

---

1 See *Heider & Simmel*: An Experimental Study of Apparent Behaviour, 243.
2 See *Turkle*: Authenticity in the Age of Digital Companions, 502.
3 See *Turkle*: Authenticity in the Age of Digital Companions, 502; on non-directive therapy, see *Rogers*: Significant Aspects of Client-Centred Therapy, 415–422.

> because he considered human understanding essential to the confidences a patient shares with a psychotherapist. Eliza could not understand the stories it was being told; it did not care about the human beings who confided in it. [...] If the software elicited trust, it was only by tricking those who used it.[4]

What is implicit here is the view that human users believed that the programme's emotionally involving replies were generated by a self that was capable of understanding or caring about its human interlocutor. If this were the case, the users could be devastated on learning that the 'connection' they had built was entirely one-sided.

The danger that human users may misunderstand the nature of relational artefacts and be harmed as a result is of course a serious one, especially in light of the fact that such tools are normally produced by profit-driven manufacturers and often made available by organisations who are under pressure to cut costs. It is important to ensure that users are not misled. However, as we will see below, we are in the early stages of understanding the issues in play. Similarly, the cognitive sciences have yet to fully understand what causes human beings to impute intentions to other beings or to objects, or to evaluate these intentions in a way that allows us to derive a feeling of trust.

With this in mind, this study focuses on one variant of the relational artefact, the robot, a type of tool which is known for its ability to elicit anthropomorphic projection, even in a pared-down form which is neither humanoid nor endowed with the capacity to use language. Defined by a recent study "as a machine that is able to physically interact with its environment and perform some sequence of behaviours, either autonomously or by remote control",[5] the robot offers a valuable point of focus for considering the question of trust.

2. Motion and the human–robot relationship

The fact that robots are able to elicit trust from humans is well established. Perhaps surprisingly, robots that are by no means human-like in their characteristics are among those most widely documented

4 *Turkle*: Authenticity in the Age of Digital Companions, 502.
5 *Kraus et al.*: Interactive Robots in Experimental Biology, 369–375.

as eliciting a trust response. In the military settings studied by Julie Carpenter, for example, bomb-diffusing robots resembling miniature industrial cranes have been celebrated as highly valued members of human teams, receiving military honours such as the purple heart and being mourned as fallen members of their cohort in cases where they are destroyed in action.[6]

Robots as colleagues and workers raise complex ethical questions which we are only beginning to address. Sven Nyholm and Jilles Smids have argued that although robots do not think in the way humans think and are unable to engage in trust-building interactions in the way human beings do,[7] they are nonetheless capable of out-performing human colleagues when measured against certain characteristics valued in collegial behaviour, such as "being reliable and trustworthy".[8] For example, a robot may have less difficulty than a human exhibiting fairness or impartiality, virtues which are hugely valued in collaborative contexts. Nyholm and Smids note that the ability of robots to perform set tasks predictably and reliably is valued by colleagues, and they suggest that the dynamics of collegial collaboration need to be distinguished from those governing emotionally driven relationships such as love and friendship. In an analysis of the ethics of robots as carers and companions, Mark Coeckelbergh has argued that while the human–robot relationship is structurally one-sided, the empathetic impulses felt by a robot's human partner are not without value in their own right;[9] this is a point to which we will return below.

This study offers an overview of recent work in the cognitive sciences that sheds light on the mechanisms by which human beings evaluate the nature, and the relational potential, of the beings and objects we encounter. We will focus specifically on human engagement with robotic movement, aiming to keep the focus on movement itself rather than a robot's additional characteristics, such as shape or its ability to use language. As we attempt to learn what it

6 See *Carpenter*: The Quiet Professional, 2013.
7 A point made by *Groom & Nass*: Can Robots Be Teammates? Benchmarks in Human–Robot Teams. Interaction Studies, 483–500, cited in *Nyholm & Smids:* Can a Robot Be a Good Colleague?, 2180.
8 *Nyholm & Smids*: Can a Robot Be a Good Colleague?, 2185.
9 See *Coeckelbergh*: Artificial Companions: Empathy and Vulnerability Mirroring in Human-Robot Relations, 1–17.

is about robots that allows human beings to trust them, we will also find ourselves asking what it is about human beings that makes us able to offer trust to robots.

Focusing on the robot's mechanical ability to move through space has analytical value because the recognition of movement has a distinctive power to provoke the human brain to impute intention. In the case of the moving geometric shapes described above, it was motion that Heider and Simmel identified as salient. Once it was set into motion, they concluded, even a two-dimensional triangle was given the perceptual status of a 'person'; it followed that "acts of persons have to be viewed in terms of motives".[10] Put simply, Heider and Simmel argued, the observers understood the simple fact of movement as something that must be accounted for.

Further, Wilma A. Bainbridge's experiments with Nico, a friendly humanoid robot encountered by participants in an office setting, demonstrated that the difference between two-dimensional movement and movement in three-dimensional space was significant.[11] In Bainbridge's studies the response of participants differed depending on whether a collaborative task performed in tandem with Nico took place in physical reality or via a video link. Study participants collaborated with Nico on manual tasks like moving books, which allowed researchers to observe how participants greeted the robot, worked alongside it and responded to instructions it delivered. Some participants worked directly with a physically present Nico, while others interacted with the robot through a live video feed. The results showed that participants of in-person collaborations were more likely to comply with unusual or nonsensical instructions delivered by Nico and rated the interactions as more positive and natural than those who collaborated via video.

In what follows, we will consider the role movement and physical presence plays in shaping how human subjects perceive robots, beginning with a brief review of how contemporary neuroscience sees the brain's capacity to analyse sensory data, and how it constructs a notion of agency. This understanding is crucial because it informs how we recognise and interact with entities that may possess—or appear to possess—agency, including robots.

10 *Herder & Simmel*: An Experimental Study of Apparent Behavior, 258.

11 See *Bainbridge et al.*: The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents, 41–52.

3. Cognition and sensory experience

How we understand the processing of sensory experience is changing rapidly, as our understanding of the brain changes. In the mid to late twentieth century, it was believed that the various functions of the brain had emerged at successive evolutionary phases, with emotion assigned to the earliest developmental layers and reason assigned to the more recent or advanced.[12] From the 1980s, scientists began to move away from this model, seeing the brain as a welter of highly adaptive networks.[13] Mounting experimental evidence has demonstrated the brain's neuroplasticity (its ability to repurpose and reorganise synaptic connections) and its capacity to redirect connections to new purposes in response to injury or environmental changes.[14]

In 2023 a team led by Evan Gordon put forward a new model, the Somato-Cognitive Action Network (SCAN) model,[15] which hypothesises a network that alternates between effector regions, which are responsible for specific motor outputs like hand or mouth movements, and inter-effector regions, which connect various parts of the motor cortex to higher cognitive areas such as the dorsal anterior cingulate cortex [dACC] and supplementary motor area [SMA].[16] This approach sees the brain's model of the body as dynamic, relying on integrated neural circuits to interpret sensory information and organise sensations into meaningful perceptions.

Central to most current approaches to the brain is the idea that it has evolved to anticipate changes in the environment, thus facilitat-

---

12 The most influential version of this view of human neural organisation, known as 'the triune brain' for its three components (the neocortex, the limbic system and the reptile brain), was articulated by Paul MacLean in work beginning in the 1940s, with his magnum opus, *The Triune Brain in Evolution*, published in 1990. In the English-speaking world, the triune brain achieved celebrity status thanks to Carl Sagan's 1977 study of human intelligence, *The Dragons of Eden*.

13 For a useful overview, see *Steffen et al.*: The Brain Is Adaptive Not Triune.

14 See *Marzola et al.*: Exploring the Role of Neuroplasticity in Development, Aging, and Neurodegeneration, Brain Sciences, 1610.

15 See *Gordon et al.*: A Somato-Cognitive Action Network Alternates with Effector Regions in Motor Cortex, 351–359.

16 See Ibid.

ing protection from danger or exploitation of opportunities.[17] Prediction error—in other words, surprise—is one of our most powerful sources of motivation, thanks to the dose of the neurotransmitter dopamine that is released when events do not align with expectation.[18]

One of the most significant researchers on emotional states, Lisa Feldman Barrett, argues that the experience of emotion is "an act of categorization, guided by embodied knowledge".[19] Focusing on finding the connection between sensory inputs and the generation of emotions in the brain, Barrett suggests that contextual cues can significantly influence how the brain interprets sensory stimuli. Integrating principles of embodiment and Bayesian inference, Barrett argues that the brain uses predictive models based on past experiences to interpret sensory inputs, categorising them into emotional states according to embodied knowledge and contextual cues. In this view, selfhood and emotional categorisation rely on the continuous modelling of the internal and external world which the brain uses to distinguish expected sensations from unexpected ones, and to develop a sense of bodily presence.[20]

Important here is the insight that our experience of the body itself is constructed through mental maps, which rely on a continuous process of interpretation and modelling of our internal and external world. This means that we 'know' our own bodies through the same process of modelling that allows us to know the world beyond the

17 See *Sol et al.*: Brain Size Predicts the Success of Mammal Species Introduced into Novel Environments, 63–71.

18 See *Schultz*: Reward Prediction Error, 369–371.

19 *Barrett*: Solving the Emotion Paradox: Categorization and the Experience of Emotion, 20. For a useful overview of how current approaches to emotion relate to the problem of brain models, see *Fernandez et al.*: Affective Experience in the Predictive Mind, 10847–10882.

20 See *Bechtel*: Representations and Cognitive Explanations, 296–306. For a useful overview of the increased interest in the social dimension of embodied cognition, including simulation as embodied practice, intercorporeality and intersubjectivity, see *Lindblom*: A Radical Reassessment of the Body in Social Cognition. Research by Vicario discusses the debate between embodied and disembodied theories of cognition, highlighting the contrasting views on the interplay between cognition and sensorimotor systems. *Vicario*: Perceiving Numbers Affects the Internal Random Movements Generator, 1–6. See also *Barsalou et al.*: Grounding Conceptual Knowledge in Modality-Specific Systems.

body.[21] The distinction we draw between self and not-self is not the result of sensing our bodies more directly than we sense the world around us.

4. Putting the focus on spatial relationships—self, other and the problem of movement through space

To grasp why motion is so important to us cognitively, it is useful to consider how we experience the body in space. Drawing on the mid-century psychologist James Gibson's concept of "affordances"[22] provides a critical vocabulary with which to explore how individuals respond adaptively to cues derived from spaces and the objects within them. It also helps us understand how individuals develop a relational and transactional framework to make sense of the stable and moving stimuli they encounter.

The work of Barbara Tversky has shed light on the centrality of spatial orientation to our interaction with the world. Tversky proposes that the brain employs a spatial framework as a universal strategy for processing all types of information, not just physical or visual inputs. The brain feeds all inputs, including abstract concepts, into spatial frameworks, which is what allows them to become tangible in our minds. Tversky argues that these spatial frameworks have an analytical capacity: for example, while the space around the body is experienced as three-dimensional, we easily reduce it to two dimensions when assessing space with navigation through it in mind.[23]

Complementing Tversky's insights, research on rodents has revolutionised our understanding of spatial cognition, above all in the discovery of the specialised neurons known as grid and place cells. In a groundbreaking series of publications in the 1970s, John O'Keefe and Jonathan Dostrovsky identified 'place cells', neurons in the rat hippocampus that activate when the animal passes through

21 See *Noel et al.*: Rapid Recalibration of Peri-Personal Space, 5089.

22 *Wit et al.*: Affordances and Neuroscience: Steps Towards a Successful Marriage, 622–629. *Gibson*: The Senses Considered as Perceptual Systems, 5.

23 See *Tversky*: Structures of Mental Spaces: How People Think About Space, 66–80.

specific locations, forming a mental map of these spaces.[24] Building on O'Keefe's work, May-Britt and Edward Moser identified a new group of cells in the rodent brain, 'grid cells', which create virtual maps by firing in grid-like patterns, with each grid cell contributing a data point to the pattern as the animal moves through space.[25] Subsequent work has suggested that human spatial experience functions similarly.[26] These neurons enable us to differentiate ourselves from other entities based on location and movement, and they are crucial for understanding social orientation and relationships.

5. Mirror neurons: The cognitive provocation of movement

Perhaps the most widely known work on how the brain perceives movement is that of Vittorio Gallese and colleagues on mirror neurons. In 1992, a team at the University of Parma studying Macaque monkeys identified a group of motor neurons located in the brain's ventral premotor area F5 that fired both when the monkeys executed actions related to reaching and grasping, and also when they observed other monkeys executing the same action.[27] In a later publication, Gallese and colleagues coined the term 'mirror neurons' to refer to these cells, suggesting that when we watch others performing an action, the same neural patterns are activated as if we were performing the action ourselves.[28]

Debate over the role of mirror neurons has given a new focus to the contrast between researchers who understand the experience of embodiment as integral to human cognition and conceptual processing, and those who understand cognition as computational and

24 See *O'Keefe & Dostrovsky*: The Hippocampus as a Spatial Map, 171.
25 See *Moser et al.:* Microstructure of a Spatial Map in the Entorhinal Cortex, 801–806. *Sargolini et al.:* Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex, 758.
26 See *Burgess, Barry & Doeller*: Evidence for Grid Cells in a Human Memory Network, 657–661.
27 See *Di Pellegrino et al.*: Understanding Motor Events, 179.
28 See *Gallese et al.*: Action Recognition in the Premotor Cortex, 604.

probabilistic.[29] The discovery of mirror neurons shifted the emphasis in a very concrete way: to build an understanding of the actions of another agent, the brain simulates these same actions in an embodied way. As Gallese and colleagues put it in a 2007 study, this is "a mandatory, nonconscious, and pre-reflexive mechanism that is not the result of a deliberate and conscious cognitive effort".[30]

The concept of mirror neurons thus challenges the idea that our cognitive and physical experiences are neatly separated; rather, the suggestion is that both our experience and our social understanding are based on the same mechanisms for internally simulating experience.

> My embodied simulation model is in fact challenging the notion that the sole account of interpersonal understanding consists in explicitly attributing to others propositional attitudes like beliefs and desires, mapped as symbolic representations. Before and below mind reading is intercorporeity as the main source of knowledge we directly gather about others.[31]

Gallese sees intersubjectivity not as mediated by the sensory cortex, where sensory stimuli are processed, but instead, by the motor cortex, where the brain generates actions.[32] To illustrate the point, he brings infant development into play: "infants develop the capacity to anticipate the goal of the observed motor acts done by others only when they become able to perform the same goal-directed motor acts themselves."[33]

The debate over the extent and nature of mirror neurons remains open. A recent survey of studies of their function concludes that while mirror-neuron brain areas contribute to low-level processing

---

29 See *Gallese & Sinigaglia*: What Is So Special About Embodied Simulation?, 512; *Chater et al.*: Probabilistic Models of Cognition: Conceptual Foundations, 288.

30 *Gallese, Eagle & Migone*: Intentional Attunement, 143.

31 *Gallese*: Mirror Neurons, Embodied Simulation, and the Neural Basis of Social Identification, 524. Gallese here draws on Merleau Ponty's concept of *intercorporéité*, "a kind of fundamental openness of the body to other bodies such that their coupling generates norms that come to affectively govern their engagement" (*Walsh*: Intercorporeity, 34). While the term is normally represented by "incorporeality" in English, it is sometimes (as here) translated as "intercorporeity".

32 See ibid., 522.

33 *Gallese, Eagle & Migone*: Intentional Attunement, 146.

of observed actions, such as distinguishing types of grip, they may not play a significant role in high-level action interpretation, such as inferring actors' intentions.[34]

According to Gallese, the sense of connection that comes from interacting with others is rooted in a recognition of shared biological processes that occurs at a physical level.

> The discovery of mirror neurons provide[s] a new empirically based notion of intersubjectivity, viewed first and foremost as intercorporeity—the mutual resonance of intentionally meaningful sensory-motor behaviours—as the main source of knowledge we directly gather about others.[35]

In this way, Gallese offers a striking vision of intersubjective solidarity: "Anytime we meet someone, we are implicitly aware of his or her similarity to us, because we literally embody it."[36]

But the existence of the android robot offers an important challenge to this assertion. Studies by Gazzola et al., Kashi & Levy-Tzedek and others have shown that the mirror neuron system responds not only to human actions but also to the actions of robots, which are by definition not living beings with whom intersubjective recognition may be shared.[37] On this view, the perception of intersubjectivity is a projection, not a recognition of empirical fact.

6. The hypothesis of the other's intention

We return here to an issue introduced earlier, the human tendency to attribute intention to things that move. This often takes the form of anthropomorphism, the human tendency to "explain nonhuman behavior as motivated by human feelings and mental states".[38]

34 See *Heyes & Catmur*: What Happened to Mirror Neurons?, 153–168.

35 *Gallese*: Mirror Neurons, Embodied Simulation, and the Neural Basis of Social Identification, 523.

36 Ibid., 524.

37 See *Kashi & Levy-Tzedek*: Smooth Leader or Sharp Follower? Playing the Mirror Game with a Robot, 147–159. *Gazzola et al.*: The Anthropomorphic Brain: The Mirror Neuron System Responds to Human and Robotic Actions, 1674–1684.

38 *Airenti*: The Cognitive Bases of Anthropomorphism, 119.

Building on the work of John Fisher in the early 1990s,[39] cognitive scientists have revised a view dominant since the nineteenth century, which saw anthropomorphism as "a category mistake, an obstacle to the advancement of knowledge, and as a psychological disposition typical of those who are immature and unenlightened, i.e. young children and 'primitive people'".[40] More recently, anthropomorphic thinking has come to be understood as a heuristic which serves "the need to make sense of the actions of other agents to reduce uncertainty concerning their behaviour".[41] A study by Epley, Waytz and Cacciopo clarifies the value of this heuristic in the context of interpreting movement: "Attributing human characteristics and motivations to nonhuman agents increases the ability to make sense of an agent's actions, reduces the uncertainty associated with an agent, and increases confidence in predictions of this agent in the future."[42] The need for speed and confidence in these assessments may be a question of evolutionary fitness, since the ability to predict and account for motion would have evolutionary value in the context of predator detection.[43]

Surprisingly, ascribing intention via anthropomorphic thinking seems to have value whether or not the human subject 'believes' that the object is genuinely capable of human mental states. A study by the psychologist Gabriella Airenti argues that humans including children have a productive ability to construct imaginative personas for objects and to use them meaningfully, even while recognising that they are fictions.[44] In other words, human beings don't need to believe that an artefact is human in order to engage meaningfully with it *as if it were human*. Indeed, Airenti suggests that in pastoral settings the effort to efface the difference between robots and humans can actually be counterproductive.

39 See *Fisher*: Disambiguating anthropomorphism, *Vidal et al*: Introducing Anthropomorphism and *Urquiza-Haas & Kotrschal*: The Mind Behind Anthropomorphic Thinking, 167 for useful discussion of the academic literature.

40 *Damiano & Dumouchel*: Anthropomorphism in Human-Robot Co-evolution, 2.

41 *Urquiza-Haas & Kotrschal*: The Mind Behind Anthropomorphic Thinking, 168.

42 *Epley*: On Seeing Human, 866.

43 See *Barrett*, Cognitive development and the understanding of animal behavior, 447–449. See also *Barrett et al.*, Accurate judgments of intention from motion cues alone.

44 See *Airenti*: The Cognitive Bases of Anthropomorphism, 122–123; *Airenti*: The Development of Anthropomorphism, 5.

"It is not simply that human-likeness is unnecessary," she warns:

> Robots that simulate them will scare people or make them think to be [sic] cheated. Instead, if we base [sic] on what we have seen already active in infants, being perceived as helpers in action will be sufficient to gain the sympathy of their human interlocutors.[45]

Humans are able to feel sympathy towards a machine which they know to be a machine, Airenti argues, but in order to trust it, there needs to be an absence of signals that reflect an intent to deceive the user about whether it is a machine or not.

It is perhaps significant that Airenti suggests that anthropomorphism is not an innate tendency in children, away from which adults strive to lead them. It seems, in fact, to be a form of behaviour which children learn from their adult carers.[46] Airenti stresses the social aspect here—both the relationship between the human observer and the artefact being observed, and the relationship between humans who cooperatively engage with an artefact. She concludes that treating an artefact as if it has human intentions is a learned relational practice, which does not necessarily reflect what the human subject believes about the nature of the artefact itself.

Finally, it is worth remembering that the attribution of human qualities to new technologies has a history. In a rhetorical study of personifications of technology reaching back to the introduction of the sewing machine and electric light, R. John Brockman explored how the human analogy has been invoked at points when new technologies were being introduced. When first brought to market, Brockman suggested, new technologies are described and even marketed using the metaphor of human personality. But as they become familiar, the metaphor fades away.[47] In the case of the sewing machine and the electric lamp, the human analogy was a tool for transition—a way to make the new technology familiar during the early phase of its availability.

With this in mind, it is worth asking what role the analogy that they are like human beings—or in some cases like other animals—plays in the case of relational artefacts such as robots. Will we still impute human characteristics to them once we have become used

45 *Airenti*: The Cognitive Bases of Anthropomorphism, 124.
46 See *Airenti*: The Development of Anthropomorphism, 7.
47 See *Brockman*: The Homunculus in the Computer?

to them? The chances are we will, given our tendency to anthropomorphise anything that moves. But like the miniature crane praised for its military service, the artefacts that prove most able to elicit trust may prove to be those which are designed not to feed the anthropomorphic imagination, but to perform in a way that is so predictable, consistent and undemanding that no one could mistake them for a human.

Bibliography

Airenti, G.: The Cognitive Bases of Anthropomorphism. From Relatedness to Empathy, in: International Journal of Social Robotics 7 (1), 2015, 117–127 https://doi.org/10.1007/s12369-014-0263-x.

Airenti, G.: The Development of Anthropomorphism in Interaction. Intersubjectivity, Imagination, and Theory of Mind, in: Frontiers in Psychology 9, 2018, 2136 https://doi.org/10.3389/fpsyg.2018.02136.

Bainbridge, Wilma A./*Hart*, Justin W./*Kim*, Elizabeth S. et al.: The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents, in: International Journal of Social Robotics 3 (1), 2011, 41–52 https://doi.org/10.1007/s12369-010-0082-7.

Barrett, H. C./*Todd*, P. M./*Miller*, G. F. et al.: Accurate Judgments of Intention from Motion Cues Alone. A Cross-Cultural Study, in: Evolution and Human Behavior 26 (4), 2005, 313–331 https://doi.org/10.1016/j.evolhumbehav.2004.08.015.

Barrett, H. C.: Cognitive Development and the Understanding of Animal Behavior, in: Elis, B. J./Bjorklund, D. F. (eds.): Origins of the Social Mind, New York 2005, 438–467.

Barrett, Lisa F.: Solving the Emotion Paradox. Categorization and the Experience of Emotion, in: Personality and Social Psychology Review 10 (1), 2006, 20–46 https://doi.org/10.1207/s15327957pspr1001_2.

Barsalou, Lawrence W./*Simmons*, W. Kyle/*Barbey*, Aron K. et al.: Grounding Conceptual Knowledge in Modality-Specific Systems, in: Trends in Cognitive Sciences 7 (2), 2003, 84–91 https://doi.org/10.1016/S1364-6613(02)00029-3.

Bechtel, William: Representations and Cognitive Explanations. Assessing the Dynamicist's Challenge in Cognitive Science, in: Cognitive Science 22 (3), 1998, 295–318 https://doi.org/10.1207/s15516709cog2203_2.

Brockmann, R. John: A Homunculus in the Computer?, in: Journal of Technical Writing and Communication 27 (2), 1997, 119–145 https://doi.org/10.2190/E7HL-A1V3-RTUY-8PE2.

Butterfill, Stephen A./*Apperly*, Ian A.: How to Construct a Minimal Theory of Mind, in: Mind & Language 28 (5), 2013, 606–637 https://doi.org/10.1111/mila.12036.

Carpenter, Julie: The Quiet Professional. An Investigation of U.S. Military Explosive Ordnance Disposal Personnel Interactions with Everyday Field Robots, unpublished dissertation, University of Washington, 2013.

Chater, Nick/*Tenenbaum*, Joshua B./*Yuille*, Alan: Probabilistic Models of Cognition. Conceptual Foundations, in: Trends in Cognitive Sciences 10 (7), 2006, 287–291 https://doi.org/10.1016/j.tics.2006.05.007.

Chemero, Anthony: Radical Embodied Cognitive Science, Cambridge 2009 https://doi.org/10.7551/mitpress/8367.001.0001.

Clark, Andy: The Dynamical Challenge, in: Cognitive Science 21 (4), 1997, 461–481 https://doi.org/10.1207/s15516709cog2104_3.

Clark, Andy: Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science, in: Behavioral and Brain Sciences 36 (3), 2013, 181–204 https://doi.org/10.1017/S0140525X12000477.

Coeckelbergh, Mark: Artificial Companions. Empathy and Vulnerability Mirroring in Human-Robot Relations, in: Studies in Ethics, Law, and Technology 4 (3), 2011 https://doi.org/10.2202/1941-6008.1126.

Colombo, Matteo/*Seriès*, Peggy: Bayes in the Brain. On Bayesian Modelling in Neuroscience, in: The British Journal for the Philosophy of Science 63 (3), 2012, 697–723 https://doi.org/10.1093/bjps/axr043.

Damiano, L./*Dumouchel*, P.: Anthropomorphism in Human-Robot Co-evolution, in: Frontiers in Psychology 9, 2018, 468 https://doi.org/10.3389/fpsyg.2018.00468.

De Wit, Matthieu M./*De Vries*, Simon/*Van Der Kamp*, John et al.: Affordances and Neuroscience. Steps towards a Successful Marriage, in: Neuroscience & Biobehavioral Reviews 80, 2017, 622–629 https://doi.org/10.1016/j.neubiorev.2017.07.008.

Di Pellegrino, G./*Fadiga*, L./*Fogassi*, L. et al.: Understanding Motor Events. A Neurophysiological Study, in: Experimental Brain Research 91 (1), 1992, 176–180 https://doi.org/10.1007/BF00230027.

Doeller, Christian F./*Barry*, Caswell/*Burgess*, Neil: Evidence for Grid Cells in a Human Memory Network, in: Nature 463.7281, 2010, 657–661 https://doi.org/10.1038/nature08704.

Epley, N./*Waytz*, A./*Cacioppo*, J. T.: On Seeing Human. A Three-Factor Theory of Anthropomorphism, in: Psychological Review 114 (4), 2007, 864–886 https://doi.org/10.1037/0033-295X.114.4.864.

Fernandez Velasco, Pablo/*Loev*, Slawa: Affective Experience in the Predictive Mind. A Review and New Integrative Account, in: Synthese 198 (11), 2021, 10847–10882 https://doi.org/10.1007/s11229-020-02755-4.

Fisher, J. A.: Disambiguating Anthropomorphism. An Interdisciplinary Review, in: Perspectives in Ethology 9, 1991, 49–85.

Freeman, Robert B.: The Senses Considered as Perceptual Systems. James J. Gibson, in: The Quarterly Review of Biology 44 (1), 1969, 104–105 https://doi.org/10.1086/406033.

Friston, Karl: The Free-Energy Principle. A Unified Brain Theory?, in: Nature Reviews Neuroscience 11 (2), 2010, 127–138 https://doi.org/10.1038/nrn2787.

Gallese, Vittorio/*Sinigaglia*, Corrado: What Is so Special about Embodied Simulation?, in: Trends in Cognitive Sciences 15 (11), 2011, 512–519 https://doi.org/10.1016/j.tics.2011.09.003.

Gallese, Vittorio/*Eagle*, Morris N./*Migone*, Paolo: Intentional Attunement. Mirror Neurons and the Neural Underpinnings of Interpersonal Relations, in: Journal of the American Psychoanalytic Association 55 (1), 2007, 131–175 https://doi.org/10.1177/00030651070550010601.

Gallese, Vittorio: Mirror Neurons, Embodied Simulation, and the Neural Basis of Social Identification, in: Psychoanalytic Dialogues 19 (5), 2009, 519–536 https://doi.org/10.1080/10481880903231910.

Gazzola, V./*Rizzolatti*, G./*Wicker*, B. et al.: The Anthropomorphic Brain. The Mirror Neuron System Responds to Human and Robotic Actions, in: NeuroImage 35 (4), 2007, 1674–1684 https://doi.org/10.1016/j.neuroimage.2007.02.003.

Gordon, Evan M./*Chauvin*, Roselyne J./*Van*, Andrew N. et al.: A Somato-Cognitive Action Network Alternates with Effector Regions in Motor Cortex, in: Nature 617.7960, 2023, 351–359 https://doi.org/10.1038/s41586-023-05964-2.

Groom, Victoria/*Nass*, Clifford: Can Robots Be Teammates? Benchmarks in Human–Robot Teams, in: Interaction Studies 8 (3), 2007, 483–500 https://doi.org/10.1075/is.8.3.10gro.

Hafting, Torkel/*Fyhn*, Marianne/*Molden*, Sturla et al.: Microstructure of a Spatial Map in the Entorhinal Cortex, in: Nature 436.7052, 2005, 801–806 https://doi.org/10.1038/nature03721.

Heider, Fritz/*Simmel*, Marianne: An Experimental Study of Apparent Behavior, in: The American Journal of Psychology 57 (2), 1944, 243–259 https://doi.org/10.2307/1416950.

Heyes, Cecilia/*Catmur*, Caroline: What Happened to Mirror Neurons?, in: Perspectives on Psychological Science 17 (1), 2022, 153–168 https://doi.org/10.1177/1745691621990638.

Kashi, Shir/*Levy-Tzedek*, Shelly: Smooth Leader or Sharp Follower? Playing the Mirror Game with a Robot, in: Restorative Neurology and Neuroscience 36 (2), 2018, 147–159 https://doi.org/10.3233/RNN-170756.

Kraišniković, Ceca/*Maass*, Wolfgang/*Legenstein*, Robert: Spike-Based Symbolic Computations on Bit Strings and Numbers, in: Frontiers in Artificial Intelligence and Applications 342, 2021, 214–234 https://doi.org/10.3233/FAIA210356.

Krause, Jens/*Winfield*, Alan F.T./*Deneubourg*, Jean-Louis: Interactive Robots in Experimental Biology, in: Trends in Ecology & Evolution 26 (7), 2011, 369–375 https://doi.org/10.1016/j.tree.2011.03.015.

Lindblom, Jessica: A Radical Reassessment of the Body in Social Cognition, in: Frontiers in Psychology 11, 2020, 987 https://doi.org/10.3389/fpsyg.2020.00987.

MacLean, Paul: The Triune Brain in Evolution, New York 1990.

Marzola, Patrícia/*Melzer*, Thayza/*Pavesi*, Eloisa et al.: Exploring the Role of Neuroplasticity in Development, Aging, and Neurodegeneration, in: Brain Sciences 13 (12), 2023, 1610 https://doi.org/10.3390/brainsci13121610.

Miller, Mark/*Clark*, Andy: Happily Entangled. Prediction, Emotion, and the Embodied Mind, in: Synthese 195 (6), 2018, 2559–2575 https://doi.org/10.1007/s11229-017-1399-7.

Noel, Jean-Paul/*Bertoni*, Tommaso/*Terrebonne*, Emily et al.: Rapid Recalibration of Peri-Personal Space. Psychophysical, Electrophysiological, and Neural Network Modeling Evidence, in: Cerebral Cortex 30 (9), 2020, 5088–5106 https://doi.org/10.1093/cercor/bhaa103.

Nyholm, Sven/*Smids*, Jilles: Can a Robot Be a Good Colleague?, in: Science and Engineering Ethics 26 (4), 2020, 2169–2188 https://doi.org/10.1007/s11948-019-00172-6.

O'Keefe, John/*Dostrovsky*, Jonathan: The Hippocampus as a Spatial Map. Preliminary Evidence from Unit Activity in the Freely-Moving Rat, in: Brain Research 34 (1), 1971, 171–175 https://doi.org/10.1016/0006-8993(71)90358-1.

Penfield, Wilder/*Boldrey*, Edwin: Somatic Motor and Sensory Representation in the Cerebral Cortex of Man as Studied by Electrical Stimulation, in: Brain 60 (4), 1937, 389–443 https://doi.org/10.1093/brain/60.4.389.

Rogers, Carl R.: Significant Aspects of Client-Centered Therapy, in: American Psychologist 1 (10), 1946, 415–422 https://doi.org/10.1037/h0060866.

Sagan, Carl: The Dragons of Eden. Speculations on the Evolution of Human Intelligence, New York 1977.

Sargolini, Francesca/*Fyhn*, Marianne/*Hafting*, Torkel et al.: Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex, in: Science 312.5774, 2006, 758–762 https://doi.org/10.1126/science.312.5774.758.

Schultz, Wolfram: Reward Prediction Error, in: Current Biology 27 (10), 2017, R369–R371 https://doi.org/10.1016/j.cub.2017.02.064.

Sol, Daniel/*Bacher*, Sven/*Reader*, Simon M./*Lefebvre*, Louis: Brain Size Predicts the Success of Mammal Species Introduced into Novel Environments, in: The American Naturalist 172.S1, 2008, 63–71 https://doi.org/10.1086/588304.

Steffen, Patrick R./*Hedges*, Dawson/*Matheson*, Rebekka: The Brain Is Adaptive Not Triune. How the Brain Responds to Threat, Challenge, and Change, in: Frontiers in Psychiatry 13, 2022, 802606 https://doi.org/10.3389/fpsyt.2022.802606.

Turkle, Sherry: Authenticity in the Age of Digital Companions, in: Interaction Studies 8 (3), 2007, 501–517.

Tversky, Barbara: Structures Of Mental Spaces. How People Think About Space, in: Environment and Behavior 35 (1), 2003, 66–80 https://doi.org/10.1177/0013916502238865.

Urquiza-Haas, E. G./*Kotrschal*, K.: The Mind Behind Anthropomorphic Thinking. Attribution of Mental States to Other Species, in: Animal Behaviour 109, 2015, 167–176 https://doi.org/10.1016/j.anbehav.2015.08.011.

Vicario, Carmelo Mario: Perceiving Numbers Affects the Internal Random Movements Generator, in: The Scientific World Journal, 2012, 1–6 https://doi.org/10.1100/2012/347068.

Vidal, Jean-Marie/*Vancassel*, Michel/*Quris*, René: Introducing Anthropomorphism, Discontinuities and Anecdotes to Question Them, in: Behavioural Processes 35 (1–3), 1995, 299–309.

Walsh, Philip J.: Intercorporeity and the first-person plural in Merleau-Ponty. Continental Philosophy Review 53, 2020, 21–47. DOI: https://doi.org/10.1007/s11007-019-09480-x

Wolpert, Daniel M./*Ghahramani*, Zoubin/*Jordan*, Michael I.: An Internal Model for Sensorimotor Integration, in: Science 269.5232, 1995, 1880–1882 https://doi.org/10.1126/science.7569931.

Ziemke, Tom: Embodiment in Cognitive Science and Robotics, in: Cangelosi, A./Asada, M. (eds.): Cognitive Robotics, Cambridge 2022, 213–230 https://doi.org/10.7551/mitpress/13780.003.0016.