Advanced Robotics Research

ADVANCED ROBOTICS RESEARCH Open Access

PERSPECTIVE OPEN ACCESS

The Future of Research in Cognitive Robotics: Foundation Models or Developmental Cognitive Models?

David Vernon

www.vernon.eu

Correspondence: David Vernon (david@vernon.eu)

Received: 15 April 2025 | **Revised:** 9 November 2025 | **Accepted:** 17 November 2025

Keywords: cognitive robotics | cognitive science | development | embodiment | enaction | foundation models | large language models | prospection | vision-language-action models

ABSTRACT

Robotics research has recently embraced foundation models as an attractive way to enhance different elements of the robotics stack and to simplify the challenge of instructing robots to perform tasks. While there is increasing evidence that this may be a very useful approach, there are also many concerns with making this the exclusive route to the future of robotics, e.g., cost, availability, trustworthiness, robustness, transparency, security, and inclusion. In this article, we will explore an alternative future for robotics research, one which eschews foundation models and deep learning with internet-scale datasets, and adopts instead an approach that has more in common with cognitive development in humans, exploiting knowledge of cognitive science while at the same time using experience in building cognitive robots to make contributions to both cognitive robotics and cognitive science. Such approaches are more aligned with the enactive stance in embodied cognitive science, in contrast with the disembodied computational functionalism of cognitivist approaches, of which foundation models are the most contemporary instance.

1 | Introduction

Foundation models [1] are increasingly being viewed as the future of robotics research [2–5]. These models are a form of generative artificial intelligence (AI) based on deep artificial neural networks that use the transformer architecture. Introduced by Vaswani et al. in 2017 [6], the transformer revolutionized deep learning. Prior to its introduction, deep learning typically used convolutional and recurrent neural networks, e.g., the pioneering AlexNet [7], VGGNet [8], and ResNet [9], among others. These so-called end-to-end networks were trained on very large datasets, differentiating them from the early deep multilayer perceptrons and long-short-term memory (LSTM) recurrent neural networks pioneered by Yann LeCun [10] and Sepp Hochreiter and Jürgen Schmidhuber [11], respectively. The transformer architecture replaced both the convolutional and LSTM layers, which require sequential training, with a multihead attention mechanism, enabling parallel training with graphic processing units (GPUs) and, consequently, allowing much larger architectures with trillions of weights, and pretraining on much larger internet-scale datasets with trillions of tokens. Foundation models originally targeted natural language processing application and provided the basis for large language models (LLMs) and related applications such as ChatGPT. However, foundation models also come in other multimodal guises, including large vision-language models (VLMs), large audio-language models (ALMs), and large visual-navigation models (VNMs). Since these models are pretrained on internet-scale data, they possess better generalization capabilities compared with traditional deep learning models that are trained on small task-specific datasets [2, 3]. Since their introduction, various techniques have been introduced to improve their already impressive performance even further, both during use through in-context learning or through querying application-specific knowledge bases, i.e., retrieval-augmented generation (RAG) [12].

There is a prevalent expectation today that foundation models will provide the same benefits in robotics as they have in natural language processing, computer vision, and image synthesis, in systems such as BERT [13], LLaMA [14], GPT-3 [15], GPT-4 [16], CLIP [17], DALL-E [18], and PaLM-E [19]. This optimism is based on the ability of multimodal foundation models to fuse heterogeneous data in compact representations. Because of their better generalization, such foundation models would provide zero-shot capabilities, i.e., the ability to perform tasks without prior examples, thereby exhibiting a generative capability. These models should enable robots to generalize across robots, environments, and tasks, and allow open vocabulary high-level task specification. By virtue of their generative character, they may allow robots to infer novel execution strategies beyond what is explicitly present in their training data.

There are two ways in which foundations models can be deployed in robotics [2]. One is by targeting components of the processing pipeline in robotics system, e.g., high-level task specification and planning, robot policy learning (e.g., language-conditioned imitation learning, and language-assisted reinforcement learning), language-image goal-conditioned value learning, and LLM-based code generation and perception. The other is by developing foundation models that are trained end-to-end on language, perception, and action data—vision-language-action models (VLA), also referred to as robot transformers—enabling robots to carry out tasks expressed in high-level abstract terms with natural language. Examples include the robotics transformer RT-1[1] [20], the RT-2 VLA model [21], and the multi-embodiment RT-X [22]. One key limitation of RT-2 and other related work is that the set of physical skills exhibited by the robot is limited to the distribution of skills observed within the robot's data [2]. Other VLA robot transformers include OpenVLA [23], CogACT [24], GROOT N1 [25], and RFM-1 [26].

However, there are still many research challenges to be overcome before achieving the goal of developing foundation models that enable robots to perform a wide range of everyday tasks safely, efficiently, and effectively, with simple natural language interfaces. These include the difficulty in obtaining internet-scale training data, dealing with the high variability of physical environments, robot morphologies, and tasks, providing safety guarantees and quantifying uncertainty, especially given the propensity of foundation models to hallucinate occasionally (i.e., generate incorrect, inconsistent, or physically infeasible output), and the difficulty in achieving real-time performance [2]. The real-time issue is exacerbated by the fact that foundation models are usually stored and executed in remote cloud-based data centers, necessitating reliable network connectivity and introducing another time limitation due to network latency. This is a significant problem in low-resource environments, such as the Global South. Other challenges include grounding for robot embodiment, adaptability to physical changes in robot embodiment, continual learning [3], decoupling robot hardware and software, robot generalization, and multimodal interaction [4]. In addition, there are also the concerns that apply to foundation models in general: the substantial cost of training such large models, with more than 100 billion parameters and training datasets in the tens of terabytes. There are also challenges in learning heterogeneous multimodal representations: it is still an open question as to whether these can be captured in simple embeddings of tokenized inputs. Finally, there is also the ever-present concern about the nascent nature of foundations models. In their seminal publication, Bommasani et al. wrote "we currently lack a clear understanding of how they work, when they fail, and what they are even capable of due to their emergent properties" [1]. There have been advances in our understanding since then, but the concern is still valid, especially given their unexpected failure modes due to their emergent operation and consequent uncertain trustworthiness.

In the following, we will explore an alternative approach grounded in developmental cognitive science, one which has been the subject of research for decades, and which offers the potential to yield insights into cognitive systems, generally, and functional advantages compared to foundation models and VLAs, in particular. The motivation for doing this is to provide some much needed balance to the current focus on foundation models and VLAs as the next generation of researchers tackle the challenges of cognitive robotics, and to remind us, as Sandini et al. have recently done [27], that there are alternative avenues of inquiry which foundation models and VLAs threaten to overshadow, and which merit our continued attention as we endeavor to advance our understanding of cognitive systems, natural and artificial. In focusing here on developmental approaches, and on development inspired by humans in particular, it is important to mention that there are other biologically inspired approaches that are also capable of producing cognitive behavior. Cognitive development is a feature of altricial species, species that are born dependent with undeveloped behaviors and skills, but with a great capacity for acquiring cognitive skills over their life-time through ontogeny. Precocial species, on the other hand, are born more self-sufficient, with well-developed behaviors, skills, and abilities as a result of their genetic make-up, i.e., their phylogeny. The challenge in cognitive robotics is to decide where on the altricial-precocial spectrum [28] to position the design of robot's cognitive architecture [29] and the degree to which one should focus on brain-based neural cognition [30–32] or cognitive behavior that results from embodied perception–action loops when interacting with the environment [33–35]. We return to the topics of embodied cognition and cognitive architecture below.

2 | Cognitive Robotics

2.1 | Definitions and Scope

There are many definitions of cognitive robotics [36], several of which emphasize the interdisciplinary nature of the field, focusing on bioinspired functionality—both human and animal[2]—across a range of skills, from sensorimotor abilities to higher-order cognitive functions and social skills. Here, cognitive science, developmental psychology, and cognitive neuroscience have had a major influence on the field. Other approaches focus more on reasoning and the use of knowledge, integrated with perception and action, reflecting the influence of early artificial intelligence (AI) on the discipline. To capture both perspectives, Cangelosi and Asada offer an inclusive definition, viz. "Cognitive robotics is the field that combines insights and methods from AI, as well as cognitive and biological sciences, to robotics" [36]. The influence of AI today extends beyond the explicit representational knowledge-based approaches to AI and, as we saw in the previous section, now includes generative AI in the form of foundation models and VLAs.

The flexible use of knowledge to support action selection and the pursuit of goals is emphasized by Sandini et al. [37], who also highlight the importance of prospection, i.e., the ability to anticipate the outcome of those actions when selecting and executing them.

The importance of social skills in a cognitive robot is reinforced by Sciutti et al. [38] who note that there is an increasing demand for social robots capable of safely interacting with people in everyday scenarios, requiring them to be able to "comprehend human behavior, predict future actions, and respond accordingly".

Significantly, there are two aspects to the goal of building a cognitive robot: one with the objective of building physical robots with the cognitive abilities of humans, and the other with the objective of gaining a better understanding of cognition (i.e., an exercise in analysis by synthesis) [39].

Generally speaking, research in cognitive robotics targets the development of relatively complete cognitive systems, integrating all the elements required for the robot to exhibit the requisite cognitive skills. This systems-oriented approach gives rise to the need for a cognitive architecture [40], a model of a cognitive system specified at several levels of abstraction, including the functional operation of individual elements and the manner in which they interact. We return to the topic of cognitive architectures in Section 5. Before that, having briefly discussed the foundations of cognitive robotics, we take a closer look at what we mean by cognition and what abilities and behavior it entails. This will provide the motivation for robotics research that eschews foundation models with internet-scale data sets, and adopts instead an approach that has more in common with cognitive development in humans.

2.2 | The Nature of Cognition

The stance we take on the nature of cognition is conditioned by the paradigm on cognitive science to which we adhere, be it the cognitivist paradigm (also known as the symbolic paradigm), the emergent paradigm, or the hybrid paradigm [41]. There are many significant differences between these three paradigms [29, 42], but the one that is particularly important here concerns the issue of embodiment. The cognitivist paradigm is grounded in computational functionalism [43], a stance that views cognition as a computational process that is independent of the physical support for the computations, i.e., the embodiment of the cognitive agent, be it is a computer or a human brain. This stance has its roots in the seminal concept of physical symbol systems introduced by Allen Newell and Herb Simon [44]. In stark contrast, in the emergent paradigm, the agent's body is an intrinsic element of cognition. In some instances, the world in which the agent is embedded is also an element of the process of cognition. In this context, we speak of embodied cognition, espoused in its purest form by enactive cognitive science [45, 46] and enaction-based cognitive robotics [47]. The cognitivist and the emergent paradigms are clearly incompatible. Hybrid approaches typically ignore this incompatibility and focus instead on models of cognition that involve both symbolic and nonsymbolic processes. These are sometimes referred to as neuro-symbolic models of cognition [48]. Embodiment entails much more than merely having a body: according to the embodied cognition thesis [49], the body and brain are both constitutive elements of the cognitive process as it interacts with its environment. Indeed, some embodied cognitive systems accomplish what can be viewed as cognitive processes exclusively by bodily means [34], sometimes referred to as morphological computing [33, 50]. For example, gaits to enable walking result from an appropriately configured body without any central controller using passive dynamics [51]. A person can position themself to catch an object by running with a speed and direction that produces a particular optical pattern (a constant velocity or tracing a straight line) as they watch the object's trajectory. These examples, along with others in [35], are instances of the replacement hypothesis of embodiment, complementing the conceptualization and constitution hypotheses of embodiment [52].

While these distinctions will be crucially important in our subsequent discussion, we can still identify common functional attributes of cognition. Cognitive agents are characterized by their ability to make inferences about the agent's world and, especially, the consequences of the agent's actions in that world. Specifically, cognitive agents have the ability to predict effects of their actions, typically when pursuing some goal, observing and assimilating the differences between its predictions and the effect of its action. It can then adapt the way it subsequently acts. This gives rise to a continuous cycle of learning and development, enhancing the agent's ability to anticipate future events and act effectively. The cycle of anticipation, assimilation, and adaptation supports—and is supported by—an on-going process of action and perception; see Figure 1. From this perspective, cognition can thus be viewed as a process by which an agent perceives its environment, learns from experience, anticipates the outcome of events, acts to pursue goals, and adapts to changing circumstances [29].

A cognitive agent should also be able to explain the reason for its actions [53], enabling it to identify problems when carrying out a task and ask for information when necessary. Thus, a cognitive agent would be able to take different perspectives on a situation, anticipating different actions and predicting their outcomes, as we mentioned above. Ideally, a cognitive agent should also have a capacity for self-reflection, often referred to as meta-cognition [54, 55].

This characterization of cognition is explicitly directed at the operation of an agent. This is a consequence of our interest in developing robots with the cognitive abilities of humans. However, it is important to be aware that there is another perspective, one that is complementary but nevertheless consistent

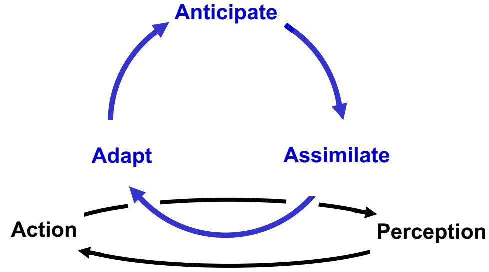

FIGURE 1 | Cognition as a cycle of anticipation, assimilation, and adaptation: embedded in a continuous process of action and perception. Reproduced with permission [29]. Copyright 2014, Massachusetts Institute of Technology. All rights reserved.

with this characterization, that views cognition as "a systemic biological function … which evolved to make survival, growth, and reproduction possible at all," a function that is present at all levels of the tree of life, "from unicellular organisms to blue whales, and every from of life in between" [56]. It is "cognition all the way down," a view of cognition that Pamela Lyon refers to as "basal cognition" [56, 57].[3]

3 | Cognitive Development

Humans develop their cognitive skills over their lifetime through ontogenesis: progressive long-term adaptation as a consequence of experience and interaction with the world. Development differs from learning in that ontogeny is the sequence of phases in which previously developed abilities are scaffolded to add to the agent's repertoire of anticipatory, prospectively controlled goal-directed actions. Ontogenesis starts with actions that require little prospection, followed by more complex actions that bring forth increasingly prospective cognitive capabilities. Thus, cognitive development both increases the system's repertoire of effective actions and extends the time-horizon of prospection, endowing an agent with the ability to anticipate the need for an action and predict the outcome of an action [58]. Prospection involves visualizing the outcome of an action as well as the predictive control of the movements entailed by action. Significantly, an action is defined more by the goal of the action than by the movements that enable the achievement of the goal. The prospective aspect of development can be accelerated by internal simulation, mentally visualizing the execution of actions and inferring the likely outcome of those actions [59, 60].

Actions are triggered by affective motives [61], as is their development. These motives don't have to be task-specific and they are selected on the basis of an intrinsic value system that constrains behavior [62]. From this perspective, the value system "mediates the saliency of environmental stimuli" [63], signaling important events and triggering the formation of goals which are then acted upon by a behavioral system. As we will see in Section 4 on autonomy, value systems are concerned not only with behavior, but also with constitutive self-organization [64–67], increasing the potential for development while maintaining the agent's autonomy.

The phased aspect of development is also manifested in the manner in which infants come to understand people's intentions and assist them in achieving their goals, taking up to 3 years before these abilities are fully formed [29, 58]. As an infant progressively acquires prospective motor skills in the first year of life, it develops an associated ability to comprehend the intentions of other agents, first by anticipating the goal of simple movements and eventually anticipating more complex goals. Around 14–18 months of age, children begin to exhibit what is referred to as instrumental helping behavior, exhibiting spontaneous helping behaviors when they see someone who is unable to achieve their goal [68]. This is a critical precursor to the subsequent development of the ability to collaborate with others, requiring shared intentions, shared goals, joint actions, and joint attention [29]. The ability to collaborate appears at about 3 years of age [69]. Since the realization of a capability for collaboration is one of the major goals of cognitive robotics [27], this provides a strong motivation for understanding of the processes by which it develops in humans. Significantly, it also provides an avenue for tackling one of the principal challenges in human–robot interaction and social robotics: the ability of a cognitive agent to form a *theory of mind* [70], also referred to as perspective taking, whereby one agent is able to take a perspective on someone else's situation and infer their intentions.

4 | Autonomy

Autonomy is an obscure concept [71] and there are several perspectives on what it means [64]. It can mean being able to operate without assistance, but it can also mean self-governance and self-regulation, and the ability of a system to determine its own goals [29, 72–75].

In robotics, we distinguish between *strength of autonomy* and *degree of autonomy* [76].[4] The former refers to the ability to of the robot to deal with uncertainty, sometimes referred to as *task entropy*; the latter refers to the degree to which a human assists the robot. We use the terms adjustable, shared, sliding, and subservient autonomy to indicate situations where the degree of autonomy can vary [29].

For biological autonomous agents, as well as bio-inspired cognitive robots, the focus of autonomy switches to the need for self-maintenance: to keep itself intact, physically and organizationally. The self-maintenance of autonomy, achieved through autonomic processes, is a crucial aspect of enactive cognitive agents [47, 77]. It also extends to the ability to adapt to unexpected damage, such as developing new gaits after a leg of a tetrapod or hexapod robot is unexpectedly damaged [78, 79]. Since cognition modulates perception and action prospectively, it is one of the primary mechanisms to maintain autonomy by anticipating precarious circumstances and the consequent need for action, thereby managing their interactions with the world in order to remain viable [80]. From this perspective, autonomy and autonomy-preserving processes are the foundation of cognition [77].

Taking these positions together, one can distinguish two forms of autonomy that are of particular relevance to cognitive robotics: behavioral autonomy and constitutive autonomy [64, 66, 81]. Behavioral autonomy is concerned with the external interactions of the agent: the extent to which it sets and achieves goals, and its robustness and adaptability in dealing with an uncertain, precarious world. Conversely, constitutive autonomy focuses on the system's internal interactions: the organizational processes that maintain its viability as an autonomous entity through on-going processes of self-construction and self-repair. While complementary, behavioral and constitutive autonomy are related because an agent can't deal with uncertainty and danger if it is not organizationally—constitutively—equipped to do so [67].

This brings us to the crux of the matter. Development is often viewed as a process of adaptation based on interaction with the environment and other agents [47, 58] and the focus is usually on behavioral autonomy. However, a consideration of constitutive autonomy suggests that the value systems that drive development are associated not just to the processes of behavioral autonomy but also to those that enable constitutive autonomy. Significantly, both forms of autonomy exhibit prospection: for effective interaction behaviors and also for effective constitutive processes, such as allostasis [82].

5 | Cognitive Architectures[5]

We noted in Section 2.1 that research in cognitive robotics usually targets the development of an entire cognitive system. This gives rise to the need for a systems-engineering perspective and a focus on integrating all the elements required for the robot to exhibit the requisite cognitive skills. The achievement of this takes the form of a cognitive architecture [40].

The term cognitive architecture originates with Allen Newell's seminal work in cognitivist cognitive science. Here, the term cognitive architecture has a very specific meaning, referring any attempt to create a unified theory of cognition [83, 84], addressing all the attributes of cognition, and doing so from a comprehensive spectrum of multidisciplinary perspectives.

Allen Newell's and John Laird's Soar architecture [85–90], John Anderson's ACT-R architecture [84, 91], and Ron Sun's CLARION[6] architecture [92, 93] are archetypal unified theories of cognition.

Building on Newell's work [83], Ron Sun defines a cognitive architecture as follows.

"A cognitive *architecture* is the overall, essential structure and process of a broadly-scoped domain-generic computational cognitive model, used for broad, multiple-level, multiple-domain analysis of cognition and behavior" [92] (emphasis in the original).

Cognitivist cognitive architectures target the aspects of cognition that are invariant over time and across contexts [94–97], using explicit knowledge to complete the cognitive model by providing the necessary context-specific information.

While the term cognitive architecture has its origins in cognitivist cognitive science, it is also used in the emergent paradigm of cognitive science to refer to the agent's phylogeny, i.e., the basis for development [29, 42].

Ron Sun's paper "Desiderata for Cognitive Architectures" identifies the attributes that a general cognitive architecture should exhibit [92]. While it recognizes the need for biological realism, it doesn't focus on the process of development. Subsequently, to complement Sun's desiderata, Vernon et al. identified ten desiderata specifically targeting developmental cognitive architectures. To a significant extent, these desiderata encapsulate the research challenges with which one is confronted when attempting to understand cognition and build a developmental cognitive robot, as well as exposing the versatility that such a robot would embody. We summarize them briefly to illustrate the breadth of the field of developmental cognitive robotics.

1. A developmental cognitive robot needs an intrinsic value system that guides action selection, directs behavior, and drives development [62, 63, 98]. This value system should encapsulate both exploratory and social motives [99, 100], as espoused in the work of pioneering developmental psychologists Jean Piaget [101] and Lev Vygotsky [102], respectively.

2. A developmental cognitive robot must be embodied [103, 104] and exhibit embodied cognition [52, 105], so that it can continually expand its action capabilities and extend its ability to anticipate the need to act and the outcome its actions.

3. A developmental cognitive robot needs mechanisms to learn the causal relationship between its actions and its experiences, i.e., its sensorimotor contingencies [106], either by motor babbling [107] or goal babbling [108, 109]. The former involves random movements while the latter targets specific motor skills. These sensorimotor contingencies form the basis for perceiving affordances [110], i.e., the use to which an object can be put by an embodied observer of the object. Affordances are dependent of the nature of the agent's embodiment.

4. A developmental cognitive robot should have the ability to categorize sensory signals without prior knowledge or external instruction, and it should be able to discern objects that exhibit biological motion, detect key facial features, and interpret facial expressions. It should have the ability to identify and locate human voices and distinguish different vocal tones.

5. A developmental cognitive robot should be able to pay attention to selected features, and it should fixate on the goal of an action, not on the component movements. The latter ability reflects the prospective nature of cognition, anticipating outcomes [108, 109, 111], as captured in the next desideratum.

6. A cognitive robot capable of development should be focused on actions that are goal-directed, triggered by motives, and guided by prospection.

7. A developmental cognitive robot must have both declarative memory (episodic and semantic) and procedural memory, capturing knowledge knowledge objects or actions [112, 113]. While semantic memory is concerned with general knowledge that is expressible in language, episodic memory [114, 115] provides the basis for prospection, i.e., the ability to simulate actions and their outcomes (see Desideratum 9) and associate actions with perceived outcomes [116].

8. A developmental cognitive robot should be capable of supervised learning, reinforcement learning, and unsupervised learning [117, 118]. Unsupervised learning and reinforcement learning support motor and goal babbling (Desideratum 3) while supervised learning and reinforcement learning support learning by imitation, a key element of development in babies [119–123].

9. A developmental cognitive robot must have a capacity for prospection, both by simulating future events [124] and simulating the execution of actions and the outcome of those actions [59, 60], in a form of mental time-travel [125].

10. The nine preceding desiderata are mainly concerned with development in the context of behavioral autonomy. However, a cognitive robot that is faithful to the emergent paradigm of cognitive science should also be constitutively autonomous, having a repertoire of processes of self-maintenance [77], effecting both homeostatic [126–129] and allostatic (i.e., adaptive and predictive) [82, 130, 131] regulation of internal parameters, driven by a value system that facilitates self-organization. The processes that support constitutive autonomy should also support behavioral autonomy, e.g., facilitating the expansion action capabilities and extending of the capacity to anticipate actions and outcomes [67].

There is, however, a concern in the community about the number of disparate cognitive architectures and the lack of agreement on what are the necessary constituents of a cognitive architecture. To address this, John Laird, Christian Lebiere, and Paul Rosenbloom have proposed a standard model of the mind [132], also known as a common model of cognition [133].

6 | Developmental Cognitive Robots

Section 1 included several examples of efforts to develop cognitive robots based on foundation models, i.e., VLA and robot transformers. But what of cognitive robots that are based on developmental cognitive science? The seminal article by Minuro Asada and colleagues on cognitive developmental robotics (CDR) [134] surveys several examples, as do Asada and Angelo Cangelosi [135] in the more recent book on cognitive robotics [136]. The iCub cognitive humanoid robot [137–140] is an iconic example of approach targeted in this perspective article, both because it is the second-most used robot in cognitive robotics research [141], after the Nao robot and ahead of the Pepper robot, and because it was specifically designed to explore and exploit developmental cognitive science. As such, it is paradigmatic example of the analysis by synthesis methodology mentioned above. The RobotCub research project [142, 143] that developed the iCub also yielded a roadmap for cognitive development in humanoid robots [58] and a prototype cognitive architecture [144]. Efforts to design a developmental cognitive architecture are still ongoing in the iCog initiative [145] at the Italian Institute of Technology, which continues to build, deploy, and support iCub robots [146]. Over the past 20 years or so, the iCub has had a major impact on cognitive robotics research in Europe, and also the USA, China, Japan, and Korea [147]. It has been used extensively to support experimental work on developmental robotics, embodied cognition, and human–robot interaction by some of the leading researchers in the field. The history of the development of the iCub is documented in [148] while the lessons learned when developing software for the iCub are discussed in [149]. Giulio Sandini explains in a 32 min documentary the motivation for designing and building the iCub, highlighting that it was the result of research activity in bioengineering whose main objective was to study humans in an endeavor to "understand how certain successful evolutionary strategies can then be transferred to new types of technologies" [150].

7 | Discussion

There is a glaring omission in the cognitive abilities listed above: the ability to communicate with humans in natural language. The reason for this is two-fold. First, many of the arguments in this article can be traced to a focus on early neonatal cognitive development in which language does not feature. Second, it is undeniable that foundation models, and LLMs in particular, offer the best natural language processing (NLP) developed to date, far surpassing classical grammar based approaches. The latency concern noted in Section 1 regarding foundation models is mitigated by the fact that NLP-based communication with humans is typically a preparatory goal-definition phase in human–robot interaction, and not part of the real-time action selection and action execution phases.

While cataloging several shortcomings of LLMs—"limited abstract reasoning or planning capabilities, limited memory, lack of autonomy, lack of human-like generalization, limited reliability and trustworthiness … producing content that is meaningless or simply false" [48], Ron Sun nevertheless argues that LLMs might play a very useful role in hybrid cognitive architectures that encapsulate dual process theory [151, 152],[7] with LLMs providing the implicit, fast, intuitive, and instinctive functionality of System 1 and the explicit, slow, deliberative functionality of System 2 being provided by symbolic processes. He demonstrates how this might be accomplished with his Clarion hybrid cognitive architecture [55, 153]. While this might well be a feasible and pragmatic way to "address the limitiations of current LLM-centered AI systems … [and] overhaul … computational cognitive architectures to better reflect advances in AI and computing technology" [48], it misses two of the central points being made in this article: that humans achieve their cognitive skills without being trained on internet-scale data, and foundation models do not help advance our understanding of cognitive science.

Furthermore, in the specific context of robotics, foundation models and VLAs address only the behavioral element of cognitive systems. As such, they have little to offer in understanding, modeling, or replicating the processes of constitutive autonomy simply because that's not how they are built or what they are built for. They are effectively highly trained autoregressive probabilistic generators of the output that is most likely to be deemed feasible in the current context, given their initial and subsequent prompts, and possibly with additional training using retrieval-augmented generation (RAG) [12]. Neither can they emulate cognitive development value-driven action selection, because they have neither intrinsic nor extrinsic values system that provide the motives for action selection.

Calling for the thorough integration of developmental robotics [134] and social robotics [154, 155], Sandini et al. make a strong case that LLMs are inherently incapable of providing the basis for collaborative cognitive robots: "generative AI cannot cooperate with humans, because it does not really understand humans and, thus, is unable to support mutual understanding" [27]. They support this statement on epistemological grounds, arguing that, since LLMs are a form of generative artificial intelligence, they focus of knowledge that is independent of an agent's personal experience and that is intrinsically disembodied, having been garnered by training on internet-scale data sets. In contradistinction, they argue, as we have done in this article, that the knowledge of a cognitive agent is fundamentally embodied and experiential, constructed by the agent as it develops through cognitive bootstrapping: a self-organizing process driven by value systems [156] (cf. Desiderata 1 and 10 above). They put it succinctly when they say that, from this cognitive perspective, "knowledge is …a process of action and being acted on", and they emphasize the importance of embodied cognition in a physical and social environment to support prospection through the construction of an internal model grounded in "personal and personalized knowledge" (cf. Desiderata 2, 3, 6, and 9).

8 | Conclusion

Cognitive robotics is a very rich field, offering many disparate challenges and research opportunities, ultimately benefiting both cognitive robotics and cognitive science, providing a means to enrich and deepen our understanding of both fields of study. It is important to emphasize that we are not suggesting that the future of research in cognitive robotics should ignore the advancement of foundation models and VLAs—as shown in Figure 2 summarizing the relative degree to which approaches based on foundation models and those based on developmental

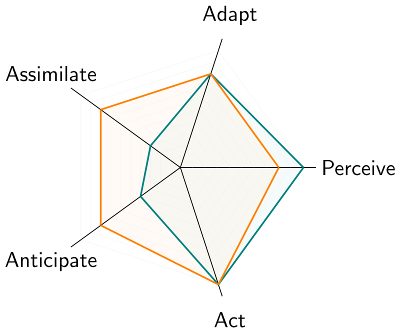

FIGURE 2 | The relative degree to which approaches based on foundation models (green) and models based on developmental cognitive science (orange) have the potential to address the five attributes of cognition shown in Figure 1, i.e., the ability to perceive the environment, learn from experience, anticipate the outcome of events, act to pursue goals, and adapt to changing circumstances [29]. Foundation models merit a high score for adaptation due to their inherent strength in generalization, but lower scores for anticipation and assimilation due to their (current) limitations in reasoning and, especially, continual learning.

cognitive science have the potential to address the five attributes of cognition in Figure 1, foundation models may well have an important role to play in contributing to the processing pipeline or as tools in HRI or cognitive bootstrapping [27]—but rather that the future of research in cognitive robotics should (re-) embrace even more strongly the enactive, embodied, prospective, developmental approaches sketched in this article.

Acknowledgments

I wish to acknowledge the insights of Prof. Giulio Sandini, with whom I have had many conversations over the past 40 years, and who had a formative influence on the views expressed in this article. These views were also informed by collaborations with Prof. Claus von Hofsten and Prof. Luciano Fadiga. I am indebted to each of you. I thank the two anonymous reviewers for their valuable remarks and suggestions on previous drafts.

Funding

The author has nothing to report.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Endnotes

[1] Strictly speaking, RT-1 is not a foundation model-based system as it was not trained on an internet-scale data set; to train the model, the authors use a (still large) dataset of over 130k real-world robotic experiences, comprising more than 700 tasks, collected using a fleet of 13 robots over a period of 17 months.

[2] Baluška and Levin [157], taking the position that cognition "refers to the total set of mechanisms and processes that underlie information acquisition, storage, processing, and use, at any level of organization," make the case that cognition is not restricted to humans and animals, and that it is also exhibited in a non-neural manner by single cells, slime molds, and plants, as well as animal tissue.

[3] Pamela Lyon's definition of cognition reads as follows.Cognition is comprised of sensory and other information-processing mechanisms an organism has for becoming familiar with, valuing, and interacting productively with features of its environment in order to meet existential needs, the most basic of which are survival/persistence, growth/thriving, and reproduction [56] (p. 416). It is significant that, as Lyon states, this co-extension of biological organization and cognition is redolent of the framework of autopoiesis (literally self-production) introduced by Humberto Maturana and Francisco Varela [158], which eventually led to the creation of the paradigm of enactive cognition [45, 46] mentioned earlier in Section 2.2.

[4] This is sometimes referred to as the *level of autonomy*, especially in human–robot interaction (HRI) [159].

[5] This section follows closely the treatment in [160]. Please refer to it for a more detailed treatment.

[6] From 2018 on, Ron Sun writes it as Clarion.

[7] The dual process or two systems (fast System 1 and slow System 2) approaches to human thought have a long history, dating back to [161]; see [162] and [151] for details.

References

1. R. Bommasani, D. A. Hudson, E. Adeli, et al., "*On the Opportunities and Risks of Foundation Models*," preprint, arXiv:2108.07258, July 12, 2022.

2. R. Firoozi, J. Tucker, S. Tian, et al., "Foundation Models in Robotics: Applications, Challenges, and the Future," *The International Journal of Robotics Research* 44 (2024): 701–739.

3. Y. Hu, Q. Xie, V. Jain, et al., "*Toward General-Purpose Robots via Foundation Models: A Survey and Meta-Analysis*," preprint, arXiv:2312.08782, October 1, 2024.

4. X. Xiao, J. Liu, Z. Wang, et al., "Robot Learning in the Era of Foundation Models: A Survey," *Neurocomputing* 638, no. 129963 (2025).

5. S. Yang, O. Nachum, Y. Du, J. Wei, P. Abbeel, and D. Schuurmans, "*Foundation Models for Decision Making: Problems, Methods, and Opportunities*," preprint, arXiv:2303.04129v1, March 7, 2023.

6. A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention Is All You Need," in Advances in Neural Information Processing Systems, Vol. 30, ed. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc, 2017), https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

7. A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems, Vol. 25, ed. F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2012).

8. K. Simonyan and A. Zisserman, "*Very Deep Convolutional Networks for Large-scale Image Recognition*," in International Conference on Learning Representations (ICLR) (2015), https://www.robots.ox.ac.uk/~vgg/publications/2015/Simonyan15/.

9. K. He, X. Zhang, S. Ren, and J. Sun, "*Deep Residual Learning for Image Recognition*," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE, 2016), 770–778, https://ieeexplore.ieee.org/document/7780459.

10. Y. LeCun, B. Boser, J. S. Denker, et al., "Back-Propagation Applied to Handwritten Zip Code Recognition," *Neural Computation* 1, no. 4 (1989): 541–551.

11. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation* 9, no. 8 (1997): 1735–1780.

12. P. Lewis, E. Perez, A. Piktus, et al., "*Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*," preprint, arXiv:2005.11401v4, April, 12, 2021.

13. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "*BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding*," preprint, arXiv:1810.04805, May 24, 2019.

14. H. Touvron, T. Lavril, G. Izacard, et al., "*LLaMA: Open and Efficient Foundation Language Models*," arXiv:2302.13971, February 27, 2023.

15. T. B. Brown, B. Mann, N. Ryder, et al., "*Language Models are Few-Shot Learners*," in Proceeding 34th Conference on Neural Information Processing Systems (NeurIPS), Vancouver, Canada (2020).

16. OpenAI, "*GPT-4 Technical Report*," preprint, arXiv: 2303.08774, March 4, 2024.

17. A. Radford, J. W. Kim, C. Hallacy, et al., "*Learning Transferable Visual Models from Natural Language Supervision*," in Proceedings of Machine Learning Research, Vol. 139, ed. M. Meila and T. Zhang (PMLR, 2021), 8748–8763, https://proceedings.mlr.press/v139/.

18. A. Ramesh, M. Pavlov, G. Goh, et al., "*Zero-shot Text-to-image Generation*," in International Conference on Machine Learning (PMLR, 2021), 8821–8831, https://proceedings.mlr.press/v202/driess23a.html.

19. D. Driess, F. Xia, M. S. M. Sajjadi, et al., "*PaLM-E: an Embodied Multimodal Language Model*," in Proceedings of the 40th International Conference on Machine Learning (2023), 8469–8488.

20. A. Brohan, N. Brown, J. Carbajal, et al., "*RT-1: Robotics Transformer for Real-World Control at Scale*," (2023).

21. A. Brohan, N. Brown, J. Carbajal, et al., *"RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control," in Proceedings of the 7th Conference on Robot Learning, Proceedings of Machine Learning Research*, Vol. 229 (2023), 2165–2183.

22. A. O'Neill, A. Rehman, A. Gupta, and etal and, "*Open X-Embodiment: Robotic Learning Datasets and RT-X Models*," preprint, arXiv:2310.08864, May 14, 2025.

23. M. J. Kim, K. Pertsch, S. Karamcheti, et al., "*OpenVLA: An Open-Source Vision-Language-Action Model*," preprint, arXiv:2406.09246, September 5, (2024).

24. Q. Li, Y. Liang, Z. Wang, et al., "*CogACT: A Foundational Vision-Language-Action Model for Synergizing Cognition and Action in Robotic Manipulation*," preprint, arXiv:2411.19650, November 29, (2024).

25. NVIDIA, GR00T N1: An Open Foundation Model for Generalist Humanoid Robots," *Technical Report* (NVIDIA, 2025), https://d1qx31qr3h6wln.cloudfront.net/publications/GR00T_1_Whitepaper.pdf.

26. A. Sohn, A. Nagabandi, C. Florensa, et al., "*Introducing RFM-1: Giving Robots Human-Like Reasoning Capabilities*," accessed 2024, https://covariant.ai/insights/introducing-rfm-1-giving-robots-human-like-reasoning-capabilities.

27. G. Sandini, A. Sciutti, and P. Morasso, "Collaborative Robots with Cognitive Capabilities for industry 4.0 and beyond," *AI* 5 (2024): 1858–1869.

28. A. Sloman and J. Chappell, "*The Altricial-Precocial Spectrum for Robots*," in IJCAI '05–19th International Joint Conference on Artificial Intelligence (2005).

29. D. Vernon, Artificial Cognitive Systems—A Primer (MIT Press, 2014).

30. J. L. Krichmar and G. M. Edelman, "Brain-Based Devices for the Study of Nervous Systems and the Development of Intelligent Machines," *Artificial Life* 11 (2005): 63–77.

31. J. L. Krichmar and G. M. Edelman, "*Principles Underlying the Construction of Brain-based Devices*," in Proceedings of AISB '06 - Adaptation in Artificial and Biological Systems, ser. Symposium on Grand Challenge 5: Architecture of Brain and Mind, ed. T. Kovacs and J. A. R. Marshall (University of Bristol, 2006), 37-42.

32. T. Hwu and J. Krichmar, Neurorobotics: Connecting the Brain, Body and Environment (MIT Press, 2022).

33. R. Pfeifer and F. Iida, "*Morphological Computation: Connecting Brain, Body, and Environment*," in AI 2006: Advances in Artificial Intelligence, 19th Australian Joint Conference on Artificial Intelligence, Vol. LNCS 3853 (Springer, 2006).

34. R. Pfeifer and J. Bongard, How the Body Shapes the Way We Think: A New View of Intelligence (MIT Press, 2007).

35. A. D. Wilson and S. Golonka, "Embodied Cognition Is Not What You Think It Is," *Frontiers in Psychology* 4, no. 58 (2013): 1–13.

36. A. Cangelosi and M. Asada, "What is Cognitive Robotics?," in Cognitive Robotics, ed. A. Cangelosi and M. Asada (MIT Press, 2022), ch. 1, 3–18.

37. G. Sandini, A. Sciutti, and D. Vernon, "Cognitive Robotics," in Encyclopedia of Robotics, ed. M. Ang, O. Khatib and B. Siciliano (Springer, 2021).

38. A. Sciutti, M. Beetz, T. Inamura, et al., "The Present and the Future of Cognitive Robotics," *IEEE Robotics and Automation Magazine* 30, no. 3 (2023): 160–163.

39. J. L. Krichmar, "Design Principles for Biologically Inspired Cognitive Architectures," *Biologically Inspired Cognitive Architectures* 1 (2012): 73–81.

40. D. Vernon, "Cognitive Architectures," Cognitive Robotics, ed. A. Cangelosi and M. Asada (MIT Press, 2022), 191–212.

41. I. Kotseruba and J. Tsotsos, "40 Years of Cognitive Architectures: Core Cognitive Abilities and Practical Applications," *Artificial Intelligence Review* 53, no. 1 (2020): 17–94.

42. D. Vernon, G. Metta, and G. Sandini, "A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents," *IEEE Transactions on Evolutionary Computation* 11, no. 2 (2007): 151–180.

43. G. Piccinini, "The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism," *Philosophy and Phenomenological Research* 81, no. 2 (2010): 269–311.

44. A. Newell and H. A. Simon, "Computer Science as Empirical Inquiry: Symbols and Search," in Communications of the Association for Computing Machinery, Vol. 19 (ACM, 1975), 113–126.

45. F. Varela, E. Thompson, and E. Rosch, The Embodied Mind (MIT Press, 1991).

46. J. Stewart, O. Gapenne, and E. A. Di Paolo, Enaction: Toward a New Paradigm for Cognitive Science (MIT Press, 2010).

47. D. Vernon, "Enaction as a Conceptual Framework for Development in Cognitive Robotics," *Paladyn Journal of Behavioral Robotics* 1, no. 2 (2010): 89–98.

48. R. Sun, "*Can a Cognitive Architecture Fundamentally Enhance LLMs? or Vice Versa?*," preprint, arXiv:2401.10444, January 19, 2024.

49. R. A. Wilson and L. Foglia, "Embodied Cognition," in The Stanford Encyclopedia of Philosophy, ed. E. N. Zalta (2011), https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=embodied-cognition&archive=spr2017.

50. H. Hauser and J. Hughes, "Morphological Computation—past, Present and Future," *Device* 2, no. 9 (2024): 1–4.

51. T. McGreer, "Passive Dynamic Walking," *International Journal of Robotics Research* 9 (1990): 62–82.

52. L. Shapiro, Embodied Cognition (Routledge, 2011).

53. R. J. Brachman, "Systems that Know What They're Doing," *IEEE Intelligent Systems* 17, no. 6 (2002): 67–71.

54. A. Sloman, "*Varieties of Affect and the Cogaff Architecture Schema*," in Proceedings of the AISB '01 Symposium on Emotion, Cognition, and Affective Computing (AISB Press, 2001), https://kar.kent.ac.uk/13629/.

55. R. Sun, "The Importance of Cognitive Architectures: An Analysis Based on CLARION," *Journal of Experimental & Theoretical Artificial Intelligence* 19, no. 2 (2007): 159–193.

56. P. Lyon, "Of What Is "Minimal Cognition" the Half-Baked Version?," *Adaptive Behavior* 28, no. 6 (2019): 407–424.

57. P. Lyon and K. Cheng, "Basal Cognition: Shifting the Center of Gravity (again)," *Animal Cognition* 26, no. 6 (2023): 1743–1750.

58. D. Vernon, C. von Hofsten, and L. Fadiga, A Roadmap for Cognitive Development in Humanoid Robots, Ser. Cognitive Systems Monographs (COSMOS), Vol. 11 (Springer, 2011).

59. G. Hesslow, "Conscious Thought as Simulation of Behaviour and Perception," *Trends in Cognitive Sciences* 6, no. 6 (2002): 242–247.

60. G. Hesslow, "The Current Status of the Simulation Theory of Cognition," *Brain Research* 1428 (2012): 71–79.

61. R. Núñez and W. J. Freeman, Reclaiming Cognition—the Primacy of Action, Intention and Emotion (Imprint Academic, 1999).

62. G. M. Edelman, Second Nature: Brain Science and Human Knowledge (Yale University Press, 2006).

63. K. E. Merrick, "A Comparative Study of Value Systems for Self-Motivated Exploration and Learning by Robots," *IEEE Transactions on Autonomous Mental Development* 2, no. 2 (2010): 119–131.

64. T. Froese, N. Virgo, and E. Izquierdo, "Autonomy: A Review and A Reappraisal," in Proceedings of the 9th European Conference on Artificial Life: Advances in Artificial, Vol. 4648, ed. F. Almeida e Costa, L. Rocha, E. Costa, I. Harvey, and A. Coutinho (Springer, 2007), 455–465, https://doi.org/10.1007/978-3-540-74913-4_46.l.

65. T. Froese and T. Ziemke, "Enactive Artificial Intelligence: Investigating the Systemic Organization of Life and Mind," *Artificial Intelligence* 173 (2009): 466–500.

66. D. Vernon, R. Lowe, S. Thill, and T. Ziemke, "Embodied Cognition and Circular Causality: On the Role of Constitutive Autonomy in the Reciprocal Coupling of Perception and Action," *Frontiers in Psychology* 6, no. 1660 (2015): 1–13.

67. D. Vernon, "Reconciling Constitutive and Behavioural Autonomy: The Challenge of Modelling Development in Enactive Cognition," *Intellectica: The Journal of the French Association for Cognitive Research* 65 (2016): 63–79.

68. F. Warneken and M. Tomasello, "The Roots of Human Altruism," *British Journal of Psychology* 100, no. 3 (2009): 455–471.

69. J. Ashley and M. Tomasello, "Cooperative Problem Solving and Teaching in Preschoolers," *Social Development* 7 (1998): 143–163.

70. A. N. Meltzoff, "Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children," *Developmental Psychology* 31 (1995): 838—850.

71. M. A. Boden, "Autonomy: What Is It?," *Bio Systems* 91 (2008): 305–308.

72. T. Ziemke, "*The 'Environmental Puppeteer' Revisited: A Connectionist Perspective on Autonomy*," in Proceedings of the 6th European Workshop on Learning Robots (1997), https://link.springer.com/chapter/10.1007/978-0-585-29605-0_20.

73. T. Ziemke, "Adaptive Behaviour in Autonomous Agents," *Presence* 7, no. 6 (1998): 564–587.

74. N. Bertschinger, E. Olbrich, N. Ay, and J. Jost, "Autonomy: An Information Theoretic Perspective," *Bio Systems* 91, no. 2 (2008): 331–345.

75. A. Seth, "Measuring Autonomy and Emergence via Granger Causality," *Artificial Life* 16, no. 2 (2010): 179–196.

76. T. B. Sheridan and W. L. Verplank, "Human and Computer Control for Undersea Teleoperators," in Technical Report (MIT Man-Machine Systems Laboratory, 1978).

77. M. H. Bickhard, "Autonomy, Function, and Representation," *Communication and Control—Artificial Intelligence* 17, no. 3-4 (2000): 111–131.

78. J. Bongard, V. Zykov, and H. Lipson, "Resilient Machines through Continuous Self-Modeling," *Science* 314, no. 5802 (2006): 1118–1121.

79. A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret, "Robots that Can Adapt like Animals," *Nature* 521, no. 7553 (2015): 503–507.

80. W. D. Christensen and C. A. Hooker, "An Interactivist-Constructivist Approach to Intelligence: Self-Directed Anticipative Learning," *Philosophical Psychology* 13, no. 1 (2000): 5–45.

81. X. Barandiaran and A. Moreno, "Adaptivity: From Metabolism to Behavior," *Adaptive Behavior* 16, no. 5 (2008): 325–344.

82. P. Sterling, "Allostasis: A Model of Predictive Regulation," *Physiology and Behaviour* 106, no. 1 (2012): 5–15.

83. A. Newell, Unified Theories of Cognition (Harvard University Press, 1990).

84. J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An Integrated Theory of the Mind," *Psychological Review* 111, no. 4 (2004): 1036–1060.

85. J. E. Laird, A. Newell, and P. Rosenbloom, "Soar: An Architecture for General Intelligence," *Artificial Intelligence* 33 (1987): 1–64.

86. P. Rosenbloom, J. E. Laird, and A. Newell, ed. The Soar Papers: Research on Integrated Intelligence (MIT Press, 1993).

87. J. F. Lehman, J. E. Laird, and P. S. Rosenbloom, "A Gentle Introduction to Soar, an Architecture for Human Cognition," in Invitation to Cognitive Science, Volume 4: Methods, Models, and Conceptual Issues, ed. S. Sternberg and D. Scarborough (MIT Press, 1998).

88. J. E. Laird, "Extending the Soar Cognitive Architecture," in Proceedings of the First Conference on Artificial General Intelligence (IOS Press, 2008), 224–235.

89. J. E. Laird, "*Towards Cognitive Robotics*," in Proceedings of the SPIE—Unmanned Systems Technology, Vol. 7332, ed. G. R. Gerhart, D. W. Gage, and C. M. Shoemaker (SPIE Press, 2009), 73–320Z–73 320Z–11.

90. J. E. Laird, The Soar Cognitive Architecture (MIT Press, 2012).

91. J. R. Anderson, "ACT: A Simple Theory of Complex Cognition," *American Psychologist* 51 (1996): 355–365.

92. R. Sun, "Desiderata for Cognitive Architectures," *Philosophical Psychology* 17, no. 3 (2004): 341–373.

93. R. Sun, "*A Tutorial on CLARION 5.0*," accessed November 2003. https://homepages.hass.rpi.edu/rsun/sun.tutorial.pdf.

94. W. D. Gray, R. M. Young, and S. S. Kirschenbaum, "Introduction to This Special Issue on Cognitive Architectures and Human-Computer Interaction," *Human-Computer Interaction* 12 (1997): 301–309.

95. P. Langley, "*An Adaptive Architecture for Physical Agents*," in IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IEEE Computer Society Press, 2005), 18–25.

96. P. Langley, J. E. Laird, and S. Rogers, "Cognitive Architectures: Research Issues and Challenges," *Cognitive Systems Research* 10, no. 2 (2009): 141–160.

97. F. E. Ritter and R. M. Young, "Introduction to This Special Issue on Using Cognitive Models to Improve Interface Design," *International Journal of Human-Computer Studies* 55 (2001): 1–14.

98. P. Oudeyer, F. Kaplan, and V. Hafner, "Intrinsic Motivation Systems for Autonomous Mental Development," *IEEE Transactions on Evolutionary Computation* 11, no. 2 (2007): 265–286.

99. J. Lindblom and T. Ziemke, "Social Situatedness of Natural and Artificial Intelligence: Vygotsky and beyond," *Adaptive Behavior* 11, no. 2 (2003): 79–96.

100. J. Lindblom, Embodied Social Cognition, Ser. Cognitive Systems Monographs (COSMOS) (Springer, 2015) 26.

101. J. Piaget, The Construction of Reality in the Child (Basic Books, 1954).

102. L. Vygotsky, Mind in Society: The Development of Higher Psychological Processes (Harvard University Press, 1978).

103. R. Chrisley and T. Ziemke, "Embodiment," in Encyclopedia of Cognitive Science (Macmillan, 2002), 1102–1108.

104. T. Ziemke, "*What's that Thing Called Embodiment?*," in Proceedings of the 25th Annual Conference of the Cognitive Science Society, ser. Lund University Cognitive Studies, ed. R. Alterman and D. Kirsh (Lawrence Erlbaum, 2003), 1134–1139.

105. M. L. Anderson, "Embodied Cognition: A Field Guide," *Artificial Intelligence* 149, no. 1 (2003): 91–130.

106. J. K. O'Regan and A. Noë, "A Sensorimotor Account of Vision and Visual Consciousness," *Behavioral and Brain Sciences* 24, no. 5 (2001): 883–917.

107. A. Cangelosi and M. Schlesinger, Developmental Robotics: From Babies to Robots (MIT Press, 2015).

108. M. Rolf, J. J. Steil, and M. Gienger, "Goal Babbling Permits Direct Learning of Inverse Kinematics," *IEEE Transactions on Autonomous Mental Development* 2, no. 3 (2010): 216–229.

109. M. Rolf and J. J. Steil, "*Goal Babbling: A New Concept for Early Sensorimotor Exploration*," in Proceedings of Humanoids 2012 Workshop on Developmental Robotics: Can Development Robotics Yield Human-like Cognitive Abilities? (IEEE, 2012). 40–43, https://scholar.google.com/citations?view_op=view_citation&hl=en&user=LqSPTpIAAAAJ&citation_for_view=LqSPTpIAAAAJ:roLk4NBRz8UC.

110. J. J. Gibson, "The Theory of Affordances," in Perceiving, Acting and Knowing: toward an Ecological Psychology, ed. R. Shaw and J. Bransford (Lawrence Erlbaum, 1977), 67–82.

111. C. von Hofsten, "An Action Perspective on Motor Development," *Trends in Cognitive Sciences* 8 (2004): 266–272.

112. R. Wood, P. Baxter, and T. Belpaeme, "A Review of Long-Term Memory in Natural and Synthetic Systems," *Adaptive Behavior* 20, no. 2 (2012): 81–103.

113. L. Squire, "Memory Systems of the Brain: A Brief History and Current Perspective," *Neurobiology of Learning and Memory* 82 (2004): 171–177.

114. E. Tulving, "Episodic and Semantic Memory," in Organization of Memory, ed. E. Tulving and W. Donaldson (Academic Press, 1972), 381–403.

115. E. Tulving, "Précis of Elements of Episodic Memory," *Behavioral and Brain Sciences* 7 (1984): 223–268.

116. D. Vernon, M. Beetz, and G. Sandini, "Prospection in Cognitive Robotics: The Case for Joint Episodic-Procedural Memory," *Frontiers in Robotics and AI* 2, no. 19 (2015): 1–14.

117. K. Doya, "What Are the Computations of the Cerebellum, the Basal Ganglia and the Cerebral Cortex?," *Neural Networks* 12 (1999): 961–974.

118. K. Doya, "Complementary Roles of Basal Ganglia and Cerebellum in Learning and Motor Control," *Current Opinion in Neurobiology* 10 (2000): 732–739.

119. A. N. Meltzoff and M. K. Moore, "Imitation of Facial and Manual Gestures by Human Neonates," *Science* 198 (1977): 75–78.

120. A. N. Meltzoff, "Explaining Facial Imitation: A Theoretical Model," *Early Development and Parenting* 6 (1997): 179–192.

121. K. Dautenhahn and A. Billard, "*Studying Robot Social Cognition Within a Developmental Psychology Framework*," in Proceedings of Eurobot 99: Third European Workshop on Advanced Mobile Robots (IEEE, 1999), 187–194, https://ieeexplore.ieee.org/xpl/conhome/6694/proceeding?sortType=vol-only-seq&isnumber=17915&pageNumber=2.

122. A. Billard, "Imitation," in the Handbook of Brain Theory and Neural Networks, ed. M. A. Arbib (MIT Press, 2002), 566–569.

123. A. N. Meltzoff, "The Elements of a Developmental Theory of Imitation," The Imitative Mind: Development, Evolution, and Brain Bases, eds. A. N. Meltzoff and W. Prinz (Cambridge University Press, 2002), 19–41.

124. M. E. P. Seligman, P. Railton, R. F. Baumeister, and C. Sripada, "Navigating into the Future or Driven by the past," *Perspectives on Psychological Science* 8, no. 2 (2013): 119–141.

125. D. L. Schacter, D. R. Addis, and R. L. Buckner, "Episodic Simulation of Future Events: Concepts, Data, and Applications," *Annals of the New York Academy of Sciences* 1124 (2008): 39–60.

126. A. R. Damasio, Looking for Spinoza: Joy, Sorrow and the Feeling Brain (Harcourt, 2003).

127. A. Morse, R. Lowe, and T. Ziemke, "*Towards an Enactive Cognitive Architecture*," in Proceedings of the First International Conference on Cognitive Systems (2008), https://www.researchgate.net/publication/237228297_Towards_an_Enactive_Cognitive_Architecture.

128. T. Ziemke and R. Lowe, "On the Role of Emotion in Embodied Cognitive Architectures: From Organisms to Robots," *Cognition and Computation* 1 (2009): 104–117.

129. A. Damasio and G. B. Carvalho, "The Nature of Feelings: Evolutionary and Neurobiological Origins," *Nature Reviews Neuroscience* 14 (2013): 143–152.

130. P. Sterling, "Principles of Allostasis: Optimal Design, Predictive Regulation, Pathophysiology and Rational Therapeutics," in Allostasis, Homeostasis, and the Costs of Adaptation, eds. J. Schulkinc (Cambridge University Press, 2004), 17–64.

131. J. Schulkin, "Social Allostasis: Anticipatory Regulation of the Internal Milieu," *Frontiers in Evolutionary Neuroscience* 2, no. 111 (2011): 1–15.

132. J. E. Laird, C. Lebiere, and P. S. Rosenbloom, "A Standard Model of the Mind: Toward a Common Computational Framework Across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics," *AI Magazine* 38, no. 4 (2017): 13–26.

133. J. Laird, C. Lebiere, P. Rosenbloom, and A. Stocco, "A Proposal to Extend the Common Model of Cognition with Metacognition," preprint, *arXiv:2506.07807v2*, June 2025.

134. M. Asada, K. Hosoda, Y. Kuniyoshi, et al., "Cognitive Developmental Robotics: A Survey," *IEEE Transactions on Autonomous Mental Development* 1, no. 1 (2009): 12–34.

135. M. Asada and A. Cangelosi, "Developmental Robotics," in Cognitive Robotics, ed. A. Cangelosi and M. Asada (MIT Press, 2022), ch. 1, 3–18.

136. A. Cangelosi and M. Asada, ed. Cognitive Robotics (MIT Press, 2022).

137. G. Sandini, G. Metta, and D. Vernon, "The iCub Cognitive Humanoid Robot: An Open-System Research Platform for Enactive Cognition," in *50 Years of AI*, Vol. LNAI 4850, ed. R. Lungarella, F. Iida, J. C. Bongard, and R. Pfeifer (Springer, 2007), 359–370.

138. N. G. Tsagarakis, G. Metta, G. Sandini, et al., "ICub – the Design and Realisation of an Open Humanoid Platform for Cognitive and Neuroscience Research," *International Journal of Advanced Robotics* 21, no. 10 (2007): 1151–1175.

139. G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The iCub Humanoid Robot: an Open Platform for Research in Embodied Cognition," in Proceedings of the Performance Metrics for Intelligent

Systems Workshop (PerMIS), ed. R. Madhavana and E. Messina (NIST Special Publication 1090, 2008), 50–56.

140. G. Metta, L. Natale, F. Nori, et al., "The iCub Humanoid Robot: An Open-Systems Platform for Research in Cognitive Development," *Neural Networks, Special Issue on Social Cognition: From Babies to Robots* 23 (2010): 1125–1134.

141. D. Ferigo, A. Parmiggiani, E. Rampone, et al., "Robot Platforms and Simulators," in Cognitive Robotics, ed. A. Cangelosi and M. Asada (MIT Press, 2022).

142. "*The RobotCub Project*," accessed November 2025, https://cordis.europa.eu/project/id/004370.

143. "*RobotCub: Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours*," accessed November 2025, http://www.robotcub.org.

144. D. Vernon, G. Metta, and G. Sandini, "*The iCub Cognitive Architecture: Interactive Development in a Humanoid Robot*," in Proceedings of IEEE International Conference on Development and Learning (ICDL) (Imperial College, 2007).

145. iCog, "*The iCub Cognitive Architecture*," accessed November 2025, https://icog.eu/.

146. iCub, "*The iCub Website*," accessed November 2025, https://icub.iit.it/.

147. iCub, "*The iCub Locations*," accessed November 2025, https://infogram.com/copy-icub-map-1h9j6qgelxw054g?live.

148. iCub, "*The iCub History*," accessed November 2025, https://icub.iit.it/web/icub/about-us/icub-history.

149. L. Natale, A. Paikan, M. Randazzo, and D. E. Domenichelli, "The Icub Software Architecture: Evolution and Lessons Learned," *Frontiers in Robotics and AI* 3 (2016): 24.

150. iCub, "*The iCub Explained*," accessed November 2025, https://www.youtube.com/watch?v=W3gIV81GYm4/.

151. K. E. Stanovich and R. F. West, "Individual Differences in Reasoning: Implications for the Rationality Debate," *Behavioral and Brain Sciences* 23 (2000): 645–665.

152. D. Kahneman, "A Perspective on Judgement and Choice—mapping Bounded Rationality," *American Psychologist* 58, no. 9 (2003): 697–720.

153. R. Sun, Anatomy of the Mind: Exploring Psychological Mechanisms and Processes with the Clarion Cognitive Architecture (Oxford University Press, 2016).

154. A. Tanevska, F. Rea, G. Sandini, L. Cañamero, and A. Sciutti, "*A cognitive architecture for socially adaptable robots*," in Proceedings of the Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob) (IEEE, 2019), 195–200, https://ieeexplore.ieee.org/document/8850688.

155. A. Tanevska, F. Rea, G. Sandini, L. Canãmero, and A. Sciutti, "A Socially Adaptable Framework for Human-Robot Interaction," *Frontiers in Robotics and AI* 7 (2020): 121.

156. R. Pfeifer, M. Lungarella, and F. Iida, "Self-Organization, Embodiment, and Biologically Inspired Robotics," *Science* 318 (2007): 1088–1093.

157. F. Baluška and M. Levin, "On Having No Head: Cognition Throughout Biological Systems," *Frontiers in Psychology* 7 (2016): 902.

158. H. R. Maturana and F. J. Varela, Autopoiesis and Cognition—the Realization of the Living, Ser. Boston Studies on the Philosophy of Science (D. Reidel Publishing Company, 1980).

159. M. A. Goodrich and A. C. Schultz, "Human–robot Interaction: A Survey," *Foundations and Trends in Human–Computer Interaction* 1, no. 3 (2007): 203–275.

160. D. Vernon, C. von Hofsten, and L. Fadiga, "Desiderata for Developmental Cognitive Architectures," *Biologically Inspired Cognitive Architectures* 18 (2016): 116–127.

161. W. James, The Principles of Psychology, Vol. 1 (Harvard University Press, 1890).

162. S. A. Sloman, "The Empirical Case for Two Systems of Reasoning," *Psychological Bulletin* 119 (1996): 3–22.