Grounding Intelligence: From Biological Minds to Artificial Cognition

Francesco Abbate[1] · Rocco Gaudenzi[2]

Abstract

In this manuscript, we propose a classification of cognitive faculties based on the current macro-approach of grounded cognition. Within this framework, we situate the classic distinction between narrow and general intelligence, offering our conceptualisation of these categories. We argue that the difficulty of assessing intellectual abilities in the artificial domain cannot be properly addressed without first determining whether these abilities are grounded or ungrounded and understanding the implications of this distinction. In fact, we maintain that in the artificial sphere, ungrounded capabilities are not necessarily synonymous with a lack of intelligence. To account for this, we propose a disembodied conception of intelligence, one that is decoupled from basic cognitive abilities as well as from constructs such as subjectivity, selfhood, and mind. For these reasons, we also suggest conceiving of ungrounded intelligence as extended. Finally, we introduce what we call the Value Grounding Problem as a conceptual test for distinguishing between grounded and ungrounded cognitive capabilities in the artificial field.

Keywords: General and Narrow Artificial Intelligence · Grounded and Ungrounded Cognition · Extended Intelligence · Artificial Self-Development

1 Introduction

In a 1991 paper entitled Intelligence without Representation, Rodney Brooks stated that the intelligence of his robots was closer to that of insects than bacteria and that he expected to achieve insect-level intelligence within two years. More than thirty years later, it is evident not only that he was mistaken but that his prediction was overly optimistic. Even today, the intelligence of robots remains far from that of the simplest forms of animal life. In this paper,

[1] Sapienza University, Department of philosophy, 00161, Rome, Italy.

[2] University of Verona, Department of human sciences, 37129, Verona, Italy.

we analyse why Brooks was wrong and propose a qualitative criterion to distinguish the abilities—both actual and potential—of different types of artificial agents (AAs).

Following Goertzel (2014, p.3), we recognise at least two very different paths in the field of artificial intelligence (AI): narrow AI and artificial general intelligence (AGI), which he proposed along with his “core AGI hypothesis”. However, we argue that this now-traditional distinction has become obsolete in light of recent progress and not sufficient to adequately describe contemporary AI developments.[1] This classification has at least two shortcomings. First, it assumes the existence of a self, an ultimate and unique referent to which certain qualities can be attributed. Second, the terms narrow and general are misleading and vague, suggesting intelligence applies only to systems classified as possessing the latter and not the former.

Artefacts such as large language models (LLMs) and multi-modal models (MMs) are being integrated into an increasing number of areas of human praxis and technical artefacts, and continuing to characterise them as entirely devoid of intelligence may be misleading. At the same time, if AGI is understood as the transformation of machines into thinking beings—entities with minds of their own in the sense proposed by Haugeland (1989)—such an achievement remains improbable, even in the near future. While it is true that these technologies automate higher cognitive faculties such as language, decision-making, problem-solving, and learning, these abilities alone do not constitute what Russell and Norvig (2016) describe as the capacity to think and act humanely. Rather, such automation pertains to the ability to think and act rationally, performing high-level reasoning tasks that require independently formulated complex strategies while remaining decoupled from properly cognitive processes. Should we accept this capacity as sufficient to define general machine intelligence? Should we instead adopt a narrower conception? And if we do, can we still meaningfully speak of intelligence? In this paper, we seek to address these questions by redefining the theoretical framework within which this analysis is situated.

We propose a categorisation of intelligence that more accurately reflects today’s technological landscape while accounting for the fundamental differences between biological and non-biological entities. Addressing the problem of AI seriously requires challenging intellective biological determinism—the assumption that intelligence is defined solely by its

[1] The set of advanced capabilities, in some cases beyond human performance, shown by models such as GPT-4 and Gemini is broad and varied and includes interleaved text-image reasoning (see MMMU benchmark), language understanding (MMLU, HELM benchmarks), coding (HumanEval benchmark), editing (EditVal benchmark), general reasoning (MMMU benchmark), abstraction and reasoning tasks (ConceptARC benchmark), visual reasoning (VCR benchmark), causal reasoning and theory of mind (BigToM benchmark), among others. See Maslej et al. (2024) for further details.

resemblance to biologically evolved capacities. We argue that the fundamental differences between biological and non-biological entities do not necessarily imply an intellectual deficiency in the latter.

To complement the narrow–general intelligence divide, we situate it within the theoretical framework of grounded cognition (Barsalou, 2008, 2010, 2015, 2020; Kiefer & Barsalou, 2013; Pecher & Zwaan, 2005). Like grounded cognition theorists, we maintain that cognition is not modular; it cannot be separated from the broad set of processes and functions that enable it, such as sensory modalities, bodily structures, and the surrounding environment—both physical and social[2] (Bersalou, 2020). Grounding, in this context, refers to the idea that cognition necessarily emerges from the interaction of these elements. While these functions are integral to biological beings, non-biological ones do not necessarily possess them, at least not in their entirety. The superior cognitive capacities of mammals, non-human primates and humans have emerged through a hierarchical evolutionary trajectory in which they developed from progressively more basic cognitive abilities shared across the animal kingdom.[3] By

[2] Among what he calls modalities, Bersalou includes external perception, characterized by the five senses, and internal perception, in which he includes proprioception, interoception, affect, reward, and introspection. In the body he includes the set of functions and apparatuses that make movement and life possible, i.e. the musculoskeletal system, the endocrine system, the immune system, the cardiovascular system, the digestive system, etc. Within the physical environment he includes the ability to interact with the outside world and the entities that populate it, such as living things, artifacts, products of nature, etc. By social environment, he finally means the capacity to have a self and agency, of being in social groups by interacting in them, of mirroring, and of having a culture. It is clear that many of these constructs are problematic, i.e. they can be understood very differently depending on the philosophical and/or cognitive science tradition one starts from. Since that of Bersalou (2020) is also a review of the different contributions made by the theorist of this and similar tradition, such as the *4E cognition* (Newen et al. 2018; Thompson, 2010) or the theorists of the *embodied cognition* (Markman & Brendl, 2005; Ziemke, 2022), we think it is an excellent starting point for those who want to explore the way these constructs are treated by that tradition, and we refer to it for further study.

[3] On this we share MacLean's (1982, 1990) theory of the tripartite brain that we believe helps clarify the evolutive trajectory of both brain and cognitive capacities in the animal kingdom. According to MacLean, the brain, while in constant interaction with the external environment, has evolved hierarchically in three superordinate phases. The first phase, evolutionarily older, coincides with the morphology and brain functionality typical of reptiles. He speaks of the reptilian brain, anatomically composed of the brainstem, the diencephalon, the thalamostriate nuclei, and the vegetative centers, including the autonomic nervous system. The activity of these areas allows functions linked to the survival of the species, guaranteeing the satisfaction of three basic motivational systems: that of nutrition, defence, and reproduction (Panksepp, 1986, 2004). The second phase of the evolutionary trajectory of the morpho-functional development of the brain involves the constitution of the limbic brain typical of mammals. Reference is made to the mammalian brain characterized by the functional activity of the brain areas responsible for emotion: the amygdala, the hippocampus, the orbitofrontal gyrus, the hypothalamus and the cingulate gyrus (Roxo et al., 2011). The mammalian brain functionally structures the motivational systems connected with interpersonal relationships (MacLean, 1982, 1990). The latter produce a tendency to socialize in order to satisfy evolutionary and adaptive goals (Liotti, 2005). Moving from the mammalian brain we arrive at the third stage of evolution through the formation of the neocortex. This phase mainly involved primates and humans and is characterised by peculiar development of the neocortex and increased functionality of the areas of the prefrontal cortex (PFC): the orbitofrontal PFC, the dorsolateral PFC and the anterior cingulate PFC (see Faw, 2003, for more). These areas represent an executive control center that plays a key role in motor control, emotion regulation, symbolisation, planning and organisation of both reptilian and limbic motivational systems to be satisfied from time to time (Faw, 2003; Panksepp, 2014). We believe that such a perspective is sufficient to justify how, in the biological domain, cognition and soma cannot be decoupled. Indeed, the emergence of increasingly complex cognition goes hand in hand with the advent of an increasingly complex nervous system, motor and sensory system. The two systems, cognition and soma, are

contrast, non-biological beings mostly follow an ungrounded approach,[4] yet they can still act in ways that, as we will argue, can in some cases be described as intelligent.

This is not to dismiss the advantages of grounded cognition but to acknowledge the growing efficacy of various artificial approaches in automating tasks that traditionally required human, thus grounded, intelligence. We contend that it is contradictory to separate problem-solving ability from intelligence, as Floridi (2017) suggests, especially when machines are granted decision-making autonomy and when such autonomy enhances task performance. This contradiction arises because intelligence is typically attributed to those who can solve complex problems effectively without relying on chance or purely mechanical processes, which are ineffective in sufficiently variable contexts. Acknowledging the decoupling of intellective and problem-solving skills is thus incompatible with certain decision-making performances achieved in the artificial domain and results from a biological determinism that considers intelligence inseparable from cognition, meaning intelligence is recognised only when it is grounded.

Since considering cognition separately from other faculties leads to a modular conception that may be misleading, we propose to describe ungrounded artificial intellectual faculties as ungrounded intelligence or intelligence without cognition, given cognition always occurs within a larger, non-modular process. We argue that ungrounded intelligence is comparable to grounded intelligence in that both can exist in narrow and general forms. However, from our perspective, ungrounded intelligence cannot be attributed to a unified self, an indivisible referent, or a mind of its own, as it lacks intrinsic motivations and, consequently, subjectivity.

In this paper, we first classify biological intelligence within the theoretical framework of grounded cognition, situating our revised conceptualisation of narrow and general intelligence

---

therefore totally interdependent and biochemically oriented systems for species adaptation in their ecological environment (Anderson, 2007).

[4] This path is currently dominated by foundation models, in which certain higher cognitive abilities, such as learning and problem-solving, are not grounded in more basic cognitive capacities, as proposed in Barsalou s (2020) model. The lack of grounding arises from the absence of a body or, in cases where a body exists—such as in robotics-powered AI, known as AI robotics (Murphy, 2019)—the inability to integrate the fundamental cognitive capacities necessary to adequately support the higher functions with which these systems are endowed. AI robotics encompasses various approaches, including deep learning, imitation learning, reinforcement learning, and the learning of tactile and sensory motion (see Ogata et al., 2022, for more). A partial exception to this is found in robots developed under the umbrella of Cognitive Robotics (CR) (De Giacomo, 1998; Aiello et al., 2001; Levesque & Lakemeyer, 2008; Cangelosi & Asada, 2022), which can be considered partially grounded, as they attempt to develop cognition based on some of the cognitive abilities outlined in the grounded model. These approaches follow a bottom-up methodology, seeking to build higher cognitive functions upon more basic ones. However, CR encompasses multiple and varied approaches, and no unified theoretical framework establishes the minimal features an AA must possess to be considered cognitively grounded. Several studies have addressed the question of what minimum capabilities an artificial body should have (see Ziemke, 2022). In this regard, we will present our own proposal in the remainder of this paper, acknowledging that artificial grounding remains fundamentally distinct from biological grounding—unless the biological body, with all its capacities, is fully re-instantiated within an artificial one.

within it (Section 2). In Section 3, we explain why intelligence in the artificial domain is largely decoupled from basic cognitive abilities but is not necessarily absent. We identify cases in which it is appropriate to speak of ungrounded intelligence and explore the extent to which this approach can be developed. In Section 4, we propose a simplified approach to testing grounded cognition in the artificial domain by adapting Harnad s (1990) Symbol Grounding Problem (SGP) to a non-anthropocentric conception of intelligence. We also explain why the principle of no external interference, originally proposed by Harnad for the SGP, must be maintained in this context. In subsection 4.1, we clarify what we mean by this form of non-interference, specifying the basic requirements needed to comply with it. Finally, in Section 5, we further develop our conceptualisation of ungrounded intelligence in light of the previous sections.

2 On biological intelligence: classifying narrow and general intelligence within the grounded cognitive framework

The term narrow intelligence is commonly used in contrast to general intelligence[5] to describe the ability of artificial systems to operate only in restricted environments or to perform predefined tasks (Kurzweil, 2005; Goertzel, 2014). At the slightest change in initial conditions, such systems fail to operate effectively. Thus, artificial systems are considered devoid of general intelligence, as they cannot autonomously adapt their behaviour to new contexts. Most of us have encountered a customer service chatbot that could not resolve our issue, requiring human intervention to complete the task. The lack of general intelligence is typically attributed to an artificial system s inability to adjust to changing conditions independently. General intelligence, by contrast, is associated with agents that learn inductively, adapting task-specific skills to novel contexts and generalising their knowledge. Goertzel (2014) correctly ascribes such abilities to humans, implicitly suggesting that general intelligence must necessarily be human-like. However, categorising a system as narrowly intelligent tells us nothing about its intelligence; it merely states that its lacks it.

We find this dichotomy problematic because it treats intelligence in a binary manner, as either present or absent. The inevitable consequence is an oscillation between these two poles in an almost dogmatic fashion.[6] Additionally, it reduces AI to a comparison with human

[5] See Goertzel (2014) for a history of the term.
[6] According to scholars such as Floridi (2017) and Floridi and Chiriatti (2020), the intelligence of LLMs can be reduced to that of refrigerators. On the contrary, Bubeck et al. (2023) have spoken of “sparks of artificial general intelligence” referring to the performance of GPT-4.

intelligence, assuming an underlying similarity. We believe both assumptions are flawed, and that AI should be assessed on different terms.

The inadequacy of the narrow–general intelligence dichotomy becomes evident when considering the animal kingdom. It is well-established that many animals exhibit intelligence, and research has even explored the cognitive abilities of plants and bacteria.[7] However, if general intelligence is defined by reference to humans, then all non-human animals must be classified as narrowly intelligent. Given the significant variation in cognitive capacities across species, the narrow–general distinction provides little meaningful insight into their abilities. The first step, therefore, should be to de-anthropise the concept of general intelligence by defining it in broader terms rather than exclusively in relation to human cognitive abilities.

Many animal species exhibit associative learning abilities that enable them to adapt to new contexts. They learn from experience and adjust their behaviour accordingly. A classic example is associative learning through classical conditioning (Pavlov, 1941) and operant conditioning (Skinner, 1938), in which animals, such as rats in Skinner s box, learn to press a lever to obtain a reward.

At the same time, why should other forms of learning—such as social learning, where individuals acquire knowledge by observing the behaviour of others, or sensory-motor learning—not be considered? Environmental variability necessitates diverse cognitive skills, including specialised forms of knowledge. These may be acquired through individual learning or transmitted evolutionarily via innate behaviours. From an intelligence standpoint, this distinction is secondary.[8] For example, the cockroach is born with an innate ability to anticipate a toad s predatory strike and flee in the optimal direction. This skill has evolved through extreme sensitivity to air movement, allowing cockroaches to detect the toad s attack 40 milliseconds before it occurs (Camhi *et al.*, 1978). Cleverly, the cockroach does not move immediately but delays its escape until the last moment, forcing the toad to commit to a trajectory, thus increasing the cockroach s chances of survival. If the cockroach fled too early, the toad could adjust its aim and capture it. Interestingly, cockroaches that lose the sensory organs responsible for air detection—the cerci—can still learn to escape effectively.[9] Other, simpler forms of learning that are shared by most species include non-associative learning, such as sensitisation and habituation, which enable dynamic behavioural adaptation.

[7] For a review see Abbate (2023).
[8] We will discuss in the next section how innate behaviour can be seen as a form of species learning, i.e. ancestrally driven, and more generally how the species-individual distinction is too rigid and should be replaced by an evolutionary conception that is both ontogenetic and phylogenetic. For a general perspective on the latter this, see also Kupiec 2019).
[9] For more see Camhi et al. (1978).

All surviving species demonstrate intelligence relative to their ecological challenges.[10] Moving across the animal kingdom, from cognitively complex species like primates and mammals to simpler life forms like unicellular organisms, we might identify a minimum or even a zero degree of intelligence, understood as relevant and motivated agency (see also Abbate, 2023). Biological cognitive abilities should therefore be situated along a continuum, ranging from a maximum in humans to a minimum in bacteria. This perspective also suggests a multi-vectorial nature of intelligence, in which different species exhibit strengths in distinct cognitive domains. However, such capacities emerge from fundamental cognitive and perceptual mechanisms.

Can we apply the same reasoning to non-biological entities? Only if appropriate distinctions are made. We argue that most existing artificial intelligence approaches should be described as ungrounded. According to Barsalou (2020), cognition in biological systems is inherently tied to the body, sensory modalities, and environment. Most artificial systems lack this grounding. However, in robotics, some AAs possess bodily structures with sensorimotor capabilities (see Footnote 4). As a result, assessing the cognitive grounding of an AA is complex. In the next section, we will propose a functional test rather than a prescriptive list of required physical and mental attributes. First, we will examine ungrounded intelligence.

3 On ungrounded intelligence or intelligence without cognition

In the biological domain, intelligence is not a discrete property but exists along a continuum of learning and adaptive capacities. We propose applying the same perspective to the artificial domain. As observed in the biological world, reference environments—except in the case of humans—are never entirely general.[11] However, this lack of generality does not imply a total absence of intelligence, as evidenced by an animal s success in its ecological niche. The main challenge in extending this approach to artificial systems lies in the fact that their operational world—what von Uexküll (1934) termed *Wirkwelt* in reference to animals—is not tailored to the survival needs of the agent and is, therefore, not inherently related to it. As a result, assessing the intellectual capacities of machines cannot follow the same criteria used for living

[10] In the words of Manning and Dawkins (2012, p. 302): "[…] animals are adapted to their ways of life", in that of von Uexküll (1934), they have an optimal relationship with their environmental niche.

[11] In fact, it is somewhat of a stretch also to consider human intelligence as a form of general intelligence. The quasi-omni-formativity of language—and, by extension, human thought—suggests that there are no inherent limits to what humans can learn. However, if we consider only the individual, excluding the surplus of intelligence generated by the ability to offload knowledge into the environment, accumulate it over time, and share it through cooperative behaviours, the limits to learning become numerous and highly variable. This is where machines, with their vastly superior capabilities in information storage, processing speed and volume, and multitasking, possess undeniable advantages. For further discussion, see Korteling et al. (2021).

beings.

If life is not involved, any definition of "environment" is arbitrary, and with it any learning or adaptation specialised capacity. However, leaving the environment completely undefined could be a mistake, because we risk assuming intelligence as something that is not situated and exists only in highly general forms. This would be an unrealistic standard, akin to Aristotle s idea of an intellect that thinks only of itself. Instead, intelligence should be assessed in limited but sufficiently variable environments (broadly construed), where intelligent behaviour can be identified based on the agent s ability to achieve its goals in unpredictable conditions, as with animals in their niche. Any environment whose dynamics is not perfectly deterministic is suitable.[12] A relevant example is the few-shot, one-shot and zero-shot learning capabilities of large language models such as GPT. These models exhibit autonomous in-context learning without requiring extensive fine-tuning. Their success demonstrates the ability to adapt dynamically to diverse linguistic contexts despite an underlying task-agnostic architecture (Brown et al., 2020). Despite being arbitrarily defined, the human-like language environment[13] that emerges from the human-machine linguistic interaction meet the variability and unpredictability criteria: its outcomes are in no way deterministically predictable, and every linguistic act is subject to such indeterminacy.

On this basis, it is reasonable to attribute intelligence to machines in specific contexts. The possibility to define a virtually unlimited number of non-deterministic environments in which AI can perform effectively suggests intelligence is in practice multiply realisable and context-dependent, as already Turing implicitly suggested (Gaudenzi, forthcoming). There is no single, pure manifestation of intelligence; rather, intelligence is a general capacity to deploy specific faculties in unpredictable settings to achieve effective results.

However, there are clear limitations to these types of AAs. These do not concern higher cognitive functions—such as generalisation, induction or rule transfer—but the lack of more fundamental capabilities that all biological beings have. Certain ungrounded approaches may succeed in expanding the range of environments in which machines can operate effectively. For instance, the foundation model approach applied to LLMs and MMs could be extended to robotics.[14] Multitasking and meta-reinforcement learning could enhance AI s effectiveness in robotics (Yu et al., 2020), provided sufficient resources—such as data and energy—are

[12] More broadly, Tegmark (2019) defines intelligence as the ability to achieve complex goals. However, in the context of AI, this definition is problematic from our perspective, as well-designed programming could establish a sufficient set of rules to accomplish such tasks without requiring genuine intelligence. Therefore, we argue that the reference environment must be non-deterministic in order to demonstrate the true autonomy of the agent in that environment.

[13] By human-like here, we simply refer to the ability to process human language for communication purposes, rather than using it as a formative organ of thought” (von Humboldt, 1999).

[14] For more on this, see "Will generative AI transform robotics?" (2024).

available for training across contexts diverse enough from each other to be considered task-agnostic, as is already the case with MMs. Nonetheless, because AAs' cognitive capacities remain ungrounded, their goals and motivations to carry out tasks do not originate from intrinsic motivations, but are externally assigned. Consequently, ungrounded AAs resemble Searle s Chinese Room (1980) and face a modified version of Harnad s (1990) Symbol Grounding Problem. In this case, the issue is not merely the referencing of symbols (their 'grounding') but the ability to ascribe value to external entities through an independent and direct evaluation, coming from the body and the modalities, of the benefits and costs of a given action. Values do not exist in isolation; they arise within an agent s interaction with its environment—both physical and social. Sucrose, for instance, "has meaning and value as food only in the milieu that the system itself brings into existence or constitutes for itself" (Thompson, 2007, p.87). In the animal kingdom, this occurs within the context of a homeostatic system that dynamically assigning it different values depending on its internal states.

Because ungrounded artificial systems lack an intrinsic mechanism for value attribution, the values they assign to objects do not exist for them but rather for external subjects.[15] This represents a crucial distinction from living organisms, which autonomously attribute values to external objects. We term this capacity environmental grounding (EG).[16] In the absence of EG, ungrounded intelligence remains a form of relevant yet non-autonomously motivated agency within non-deterministic environments. Accordingly, we define the zero degree of ungrounded intelligence as relevant (non-autonomously motivated) agency in non-deterministic environments.

4 Grounding AI

Inspired by Harnad (1990), we refer to the problem of environmental grounding as the *Value Grounding Problem* (VGP), since it concerns the attribution of values to things without necessarily making use of symbols that stand for them—indeed several animals do not possess symbolic capacities and still are adapted to their environmental niche. We argue that

[15] In this sense, approaches such as that of vector semantics (Mollo, 2024) that propose ways to solve the grounding problem, arguably only circumvent it since the meaning of concepts remains entirely human, and it is only within this horizon that it is possible for the machine to organise the same concepts in space so as they take on meanings in relation to each other.

[16] This is our conceptual proposal for cognitive grounding in the artificial domain. Replicating the full set of bodily capacities of living beings would, in effect, mean replicating a biological body within an artificial one. Instead, from our perspective, the goal is to functionally identify the basic capacities that such a body enables. We propose using the term environmental grounding in this context, referring to the ability to ascribe values for oneself to external objects.

overcoming the VGP is essential for establishing grounded cognitive abilities in artificial systems and is more fundamental than the Symbol Grounding Problem. The VGP can be formulated as follows: How can an agent s values be intrinsic to it, that is, exist without the intervention of third parties transferring those values from the outside?

Before addressing this question, we must clarify the nature of the problem. While human language is learned rather than genetically transmitted, some may argue that an insect s motivations and values are not truly its own, as they are encoded evolutionarily and passed down genetically. In this view, the principle of non-interference would not be respected in the biological domain. Since the internal value system of non-human animals is not directly observable, it is inferred from behaviour.[17] For instance, a cockroach instinctively fears a toad as a predator, treating it as an existential threat from birth. Many other animals exhibit similarly innate avoidance behaviours.[18] This would suggest that an organism s value system is evolutionarily determined and not acquired through individual learning. However, a strict distinction between evolution and learning in animal behaviour may be misleading in this case. Evolution is shaped by ancestral behaviours, and organisms share bodily structures that enable the intergenerational transmission of the value system. Thus, while values may be genetically inherited, they remain intrinsic to the individual, as they are processed and enacted through the organism s own bodily interactions with the environment. The cockroach does not merely inherit the idea of the toad as a predator; it possesses the sensory mechanisms to detect its movements and react accordingly. For this reason, an organism s particular body is the medium through which not only given value system and goals are actually realised, but also through which these are transformed: individual learning thus contributes to shape what is evolutionary transmitted. We can describe this process as onto-phylogenetic grounding, in which the relationship between individual and species is one of co-determination. The species does not merely dictate behaviour; individual organisms also influence evolutionary trajectories.[19] By analogy, we infer that the same applies to value systems: they emerge dynamically from bodily interactions with the environment, rather than solely being imposed from evolution. In these interactions, what plays a crucial role are the endocrine system, the

[17] Given that animal behaviour is largely instinctual, we believe that it is not an oversimplification to treat values, motivations, and behaviours in the same way inferring one from the other. For humans, as we will argue, this would be an oversimplification, and it will be necessary to distinguish between the different concepts.

[18] It also involves complex behaviours, not just attack-avoid. A good example is the courtship behaviour of sage grouse, which turns out to be largely, if not entirely, genetically encoded and transmitted (see Wiley, 1973).

[19] Hansell (2000) provides compelling evidence for this interaction by studying the nest-building behaviour of swallows and martins. Despite sharing a common ancestor—identified by a molecular phylogeny using nuclear DNA—that used a single tunnelling technique, these birds evolved a diverse range of nesting strategies adapted to different ecological niches. Such adaptations illustrate how behavioural flexibility at the individual level can influence species-wide evolutionary patterns. See Hansell (2000) for details.

nervous system (where present), and other morphological features,[20] as well as the environmental feedback, both physical and social.

Thus, what guarantees value grounding in biological systems is not the absence of external influence, as the original idea of symbolic grounding contends, but the presence of an indivisible cognitive body capable of directly testing and simulating transmitted values. This is evident in human development, where certain behaviours—such as crying, sucking reflexes, and facial recognition—are innate. Infants exhibit a fundamental orientation towards other humans, seeking out faces (Johnson et al., 1991), suggesting that human cognition is shaped from birth through bodily and social interactions. These innate tendencies form the foundation for what Gallese (2007) describes as *intercorporeality*—the shared *embodied representations* that develop through emotional proximity to caregivers (Ammaniti & Gallese, 2014). Such representations are not transferred propositionally but are established through bodily experiences, shaping the subject s value system. This means that the transmission of beliefs and desires can go through the body, i.e. by determining internal states associated with specific external events or objects, which are conveyed through a variety of nonverbal communications[21] (see Schore, 2021; Feldman et al., 2010; Sieratzki & Woll, 1996, for more).

4.1 From values to motivations

As just discussed, in the animal kingdom, values and motivations are closely related. This is also evident in computational reinforcement learning (CRL) approaches (Sutton & Barto, 2014), which attempt to induce intrinsic motivations in machines by assigning rewards based on environmental interactions. In these models, the system learns to assign values to objects or actions according to its exploration of the environment. CRL approaches may be considered as capable of satisfying the VGP, since they generate machine-derived values autonomously. Indeed for many, this approach solves the problem of autonomous value assignment (see Odeyer & Kaplan, 2007). We argue that this is not the case, as the core

[20] Obviously, the cockroach fears the toad because it is vulnerable to it and serves as its prey. The morphological characteristics of its body are, therefore, a crucial factor in determining the negative value attributed to the toad. At the same time, it is precisely due to this negative value that the species has evolved specific escape behaviours, gradually specialising in a set of functions that enable the precise timing needed to evade the toad's tongue.

[21] We believe that this non-propositional form of value formation, which is given by *intercoporeality*, i.e. by one's own body and its embeddedness in a social context from which values can be shared for and through it, is an indispensable element for an adequate overcoming of the SGP and in this sense we conceive the VGP as more fundamental that the SGP. By the expression "an adequate overcoming of the SGP" we mean the possibility of determining from within a complex and multifunctional (see Jakobson, 1960) form of language which, if truly grounded, cannot do without intrinsic values and motivations. Indeed, in the absence of an adequate value grounding, language development can only be extremely simple, as is the case with Luc Steels's (2001, 2015) linguistic games in which the linguistic form is limited to the act of naming and the possibility of giving or receiving formal instructions.

issue of the Value Grounding Problem (VGP) is not merely whether machines can assign values, but rather what drives this process. Ultimately, we must ask: Who assigns the values, and what makes an agent a subject? The answer lies in the ability to acquire its own experience. The distinction between learning and experience is crucial: experience is always subjective and embodied, whereas learning is not necessarily so. One key implication is that experience cannot be transferred from one subject to another; that is, unlike learning, it is not portable.

It is important to emphasise the pivotal role of the body in this process. Consider the example of encountering a wild animal. The situation in which the animal is tied up will be vastly different from one in which it is not. Similarly, the presence or absence of its tamer will further influence our perception of risk. Each of these scenarios requires evaluation based on prior assumptions, inferring the likelihood of the animal attacking and considering the possible consequences of such an event. By extending this principle to the totality of sensations, we can view perceptions as Bayesian inferences—the best possible assumption we can make about a sensory datum (Seth, 2021). However, such an inference can only be considered grounded if it pertains directly to the subject making it—that is, if it occurs in the first person. For this reason, perception is inextricably linked to the body and the internal somatic states that any given sensation generates. For Ansermet and Magistretti (2012), the subject’s internal reality, and thus, we add, the subject itself, emerges from the tension between somatic states and the perceptions/representations associated with them. For the cognitive neuroscientist Anil Seth (2021): “[…] experiences of being you, or of being me, emerge from the way the brain predicts and controls the internal states of the body”. The process though which this control occurs is called interoception (Craig, 2002; 2009; Chritchley, 2013).

The ability to exercise such control arises only from the capacity to predict—i.e. infer—the somatic states associated with a given event. This process occurs through the inscription of psychic traces in the brain, linking certain physical states to prior experiences. When a similar situation arises, these traces can be reactivated, allowing for an adaptive response. Mental plasticity enables the formation of new neural connections through mechanisms such as Hebbian learning. Indeed, contemporary neuroscience defines experience as a synchronous and spatial association of a set of neurons whose reactivation reproduces the experience” (Ansermet & Magistretti, 2012, p. 122). It is due to neuronal plasticity that the subject s representations—what one may call my representations”—are never merely the passive result of a neutral sensation. Rather, they are shaped by the somatic state associated with that sensation, through which it ceases to be neutral and becomes somatically attuned—that is, a perception imbued with intrinsic value. It is, therefore, the set of determinations of psychic traces through experience that defines the subject s psychological capacities—what we might describe, following Ingold (1999), as a process of *en-mind-men*t.

Since this process is characterised by the inscription of traces within the nervous system—and since these traces are nothing but the result of modes of somatic activation, that is, bodily sensations—it follows that the process of *en-mind-ment* can only occur alongside, or at most in parallel with, that of *embodiment*. Both require an underlying plasticity—the capacity to change in response to external stimuli and to retain traces of those changes. In this sense, we conceive of the mind as a process emerging from a centre capable of inscribing bodily states within itself, literally marking them internally,[22] and consequently anticipating and regulating them. This perspective aligns with the insights of cybernetics pioneers Conant and Ashby (1970), who asserted that “every good regulator of a system must be a model of that system.” By this reasoning, the mind functions as a model of the body. Consequently, its capabilities and complexities mirror those of the body itself.

Thus, the solution to the VGP is as follows: The attribution of values can only be intrinsic to a subject – and therefore in compliance with the principle of non-interference - when it emerges from its own psychic or internal reality, capable of influencing representations in a way that is uniquely its own and never anyone else's. This internal reality and the motivations derived from it must be intrinsic, while values need not, strictly speaking, satisfy this requirement. Traditional psychology defines intrinsic motivation as deriving from internal satisfaction, whereas extrinsic motivation is driven by external rewards (Ryan & Deci, 2000).[23] If the action is performed with a view to an extrinsic reward, such as studying to get a good job, then the associated motivation is considered extrinsic. If, on the other hand, it is done for the sheer pleasure of doing it, it is considered intrinsic. We find this misleading and propose a different criterion: an action is intrinsically motivated if it arises from the subject s internal reality, regardless of whether it serves external goals or expectations. It is the subject who, by its own determination, decides to study with a particular goal in mind—whether it is to secure a good job, meet parental expectations or simply for the pleasure of learning (the latter being undoubtedly more satisfying). By contrast, studying can also be the result of external coercion,

[22] In this context, Antonio Damasio (1996) spoke about the somatic marker hypothesis. According to his conceptualisation, somatic markers are neural agglomerations that define the ways in which somatic-emotional mapping occurs to signify experience. They arise from the integration between environmental conditions, the body's response (its visceral processes and its bodily representation localised in the somatosensory and insular areas of the brain), and higher-level cognitive processes linked to the functioning of the ventromedial Prefrontal cortex (PFC) (Damasio, 1994, 1996). The somatic experiences, based on environmental stimuli, determine pleasant or unpleasant sensations; at this point they can be stored in the ventromedial PFC, which in turn, involving a cortico-functional pattern of activation, elaborates a presumably adaptive response, aimed at satisfying a motivational system active at that moment.

[23] For a review of this see Oudeyer and Kaplan (2007) which, apart intrinsic and extrinsic motivations distinguish between internal and external, homeostatic and heterostatic, fixed and adaptive motivations.

such as parents forcing their child to study against his or her will. Only in this case the motivation can be said to be extrinsic.

The possibility of having an experience that is mine depends on the complex interdependence between the peripheral and central areas of the body. In this sense, efforts to develop homeostatic and sentient robotics appear entirely reasonable (see Odeyer & Kaplan, 2007, for a review; see also Man & Damasio, 2019; Guerrero-Rosado & Verschure, 2021) and aligned with the challenges we have outlined. However, even if we were to grant that such systems could generate internal somatic states, despite the clear structural differences they maintain in comparison to biological systems, the question remains as to how these states could be translated into psychic traces capable of preserving them.

We are sceptical that sequential (i.e. classical) computational systems can fulfil this function. A good candidate for addressing the computational challenges associated with the Value Grounding Problem (VGP) at the central level might be the neuromorphic computing architecture. Unlike conventional digital systems, neuromorphic architectures utilise electrical spikes—unprocessed analogue inputs—rather than digital signals. This approach offers several key advantages, including the ability of individual nodes to receive inputs from and send outputs to multiple other nodes simultaneously, three-dimensional interconnections that enhance spatial density and computational efficiency compared to two-dimensional layouts, and significantly lower energy consumption, as processing occurs selectively only when required (see Marković et al., 2020, for more). However, this technology is still in its infancy, with little to no integration with homeostatic and soft robotics. As a result, it remains difficult to assess how far this approach can advance from a cognitive grounding perspective. Nevertheless, given its intrinsic characteristics, we believe that neuromorphic computing—particularly when integrated with homeostatic and soft robotics—represents the most promising approach to the problem to date. We suggest that future research should focus more explicitly on the integration of these approaches.

5 Complementing ungrounded intelligence

Given our findings so far, our concept of ungrounded intelligence must be further refined. We argue that ungrounded intelligence is realisable in multiple ways, potentially emerging from different material substrates, and fundamentally existing in the absence of both a self and a mind. Here, we understand the mind as one of the possible cognitive functions capable of dynamically integrating other cognitive functions through both conscious and unconscious processes. Ungrounded intelligence, far more than the mind,[24] should therefore be considered extended—that is, not individually contained, not necessarily embodied, and not attributable to an ultimate subject or, consequently, to a self (the former being a necessary condition for the latter).

In this sense, intelligence emerges as the outcome of a socio-technical process, manifesting as the ability to solve tasks that would otherwise require the application of the human mind. It is not merely a question of achievement or performance—in the sense that if a person behaved like this, he would be called intelligent" (McCarthy et al., 2006)—but rather the fact that such performance would be impossible without the development of certain higher cognitive functions, such as autonomous decision-making and learning. This corresponds to what we may call successful adaptive functioning in relation to assigned goals. However, ungrounded intelligence does not confer true autonomy. While these systems may exhibit autonomy in solving externally defined problems, this remains a subset of genuine agency, as it is not intrinsic to the system itself. Such machines cannot, therefore, aspire to true agency, as agentivity does not ultimately reside in anything of their own.[25] It cannot be attributed to an own mind nor to an ultimate subject. We contend that this absence of intrinsic agency represents the fundamental limitation—and ultimate dead end—of an ungrounded approach.

Conclusion

We began with the observation that a simple characterisation of AI as narrow or general is inadequate for describing current developments in the artificial domain. Moreover, this classification limits our broader understanding of intelligence and cognition, as it prevents AI from being informative about them. This binary, all-or-nothing conception of intelligence implies an underlying biological determinism, if not an anthropocentric bias. To address this, we proposed situating the analysis of any agent s intellectual capacities within the theoretical framework of grounded cognition (Barsalou, 2008, 2010, 2015, 2020; Kiefer & Barsalou, 2013; Pecher & Zwaan, 2005). Within this framework, we introduced a classification of intelligence as grounded or ungrounded, incorporating the narrow–general distinction within it. As a result, we defined four categories: grounded general intelligence, grounded narrow intelligence, ungrounded general intelligence, and ungrounded narrow intelligence.

[24] We refer to Clark and Chalmers' (1998) extended mind theory.

[25] Here applies the same principle that Harnad (2018) proposed in the case of the semantic interpretation of linguistic statements and according to which it is "always mediated by an external interpreter". Taking the example previously used on sucrose value, this means that the answer to the question "why sucrose is good" makes sense only according to the human value system.

In this context, we have examined why and in what cases neither narrow nor ungrounded intelligence should necessarily be equated with a lack of intelligence. Since most successful AI approaches, such as machine learning, deep learning models, and AI robotics, are ungrounded, we argue that assessing their intellectual capabilities within this theoretical framework allows for a clearer understanding of both their limitations and potential developments. We have discussed how, although intelligence benefits from its co-existence with a biological body, it can manifest outside of it as the ability to learn, remember, and devise autonomous strategies on behalf of others. The fact that intelligence in artificial systems is tied to specific tasks does not necessarily imply its absence—provided that the environment in which these tasks are performed is sufficiently complex and unpredictable, such that deterministic prediction of all its events cannot be defined *a priori*. In such cases, the behavioural sets required to carry out assigned tasks cannot be fully predefined by rigid rules or mechanistic programming. This is essentially why we do not believe it is correct to deny intelligence to machines capable of exceptional performance in domains such as linguistic and text-image reasoning, as demonstrated by today s multimodal models (MMs), whose abilities in many cases match or exceed human-level performance. Within this framework, we have understood intelligence as something potentially disembodied and multi-vectorial.

However, we maintain that this approach is ultimately decoupled from fundamental cognitive aspects and, therefore, from what we call the mind and subjectivity—the ultimate referent of one s goals and motivations. These elements, we argue, will always have to be sought elsewhere, rather than in specific artificial agents. This is not to categorically exclude the possibility that AI could develop effectively within the grounded domain. However, this depends on the physical capabilities that can be developed in the artificial field, which remain both far removed from and fundamentally different from those of biological organisms. The biological body possesses multiple capacities, almost all of which contribute to cognitive processes. Beyond the technical limitations that affect the possibility of transferring these properties to an artificial body, fully replicating them would amount to re-instantiating the biological body within an artificial one. To address this challenge, we proposed a simplified approach by modifying and transforming Harnad s Symbol Grounding Problem (SGP) into what we term the Value Grounding Problem (VGP).

We argue that the grounding abilities an artificial agent should possess concern the ability to assign values to external objects before assigning symbols to them. This explains why such a process must be both embodied and plastic, emerging from a dynamic interaction between peripheral and central components of the artificial body. Furthermore, we examined why we believe that the principle of non-interference in value attribution applies to the biological domain, discussing why intergenerational transmission in biological systems does not constitute a violation of this principle, but rather a necessary mechanism for the emergence of intrinsic agency and cognitive grounding.

References

Abbate, F. (2023). Natural and artificial intelligence: a comparative analysis of cognitive aspects. Minds and Machines, 33(4), 791-815.

Aiello, L. C., Nardi, D., & Pirri, F. (2012). Case studies in cognitive robotics. In *Human and Machine Perception 3: Thinking, Deciding, and Acting* (pp. 167-181). Boston, MA: Springer US

Ammaniti, M., & Gallese, V. (2014). *La nascita della intersoggettività. Lo sviluppo del sé tra psicodinamica e neurobiologia*. Raffaello Cortina Editore.

Anderson, M. L. (2007). Massive redeployment, exaptation, and the functional integration of cognitive operations. Synthese, 159(3), 329-345.

Ansermet, F., & Magistretti, P. (2012). *Gli enigmi del piacere*. Bollati Boringhieri

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, *59*(1), 617-645. DOI: https://doi. org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in cognitive science*, *2*(4), 716-724. https://doi.org/10.1111/j.1756-8765.2010.01115.x

Barsalou, L. W. (2015). Situated conceptualization: Theory and applications. In *Perceptual and emotional embodiment* (pp. 19-45). Routledge.

Barsalou, L. W. (2020). Challenges and opportunities for grounding cognition. Journal of Cognition, 3(1)

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901

Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, *47*(1-3), 139-159

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M.T., & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*. https://doi.org/10.48550/arXiv.2303.12712

Camhi, J. M., Tom, W., & Volman, S. (1978). The escape behavior of the cockroach Periplaneta americana: II. Detection of natural predators by air displacement. *Journal of comparative physiology*, *128*, 203-212.

Cangelosi, A., & Asada, M. (Eds.). (2022). *Cognitive robotics*. MIT Press

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, *58*(1), 7-19.

Conant, R. C., & Ross Ashby, W. (1970). Every good regulator of a system must be a model of that system. *International journal of systems science*, *1*(2), 89-97.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature reviews neuroscience*, *3*(8), 655-666. https://doi.org/10.1038/nrn894

Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature reviews neuroscience*, *10*(1), 59-70.

Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron*, *77*(4), 624-638.

Damasio, A., (1994). Descartes' error: emotion, reason and the human brain, New York, NY: Benguin Books.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of prefrontal cortex. Philosophical Transactions of the Royal Society of London, 351B, 1413-1420. https://doi.org/10.1098/rstb.1996.0125

De Giacomo, G. (1998). Cognitive robotics. In *Proc. AAAI'98 Fall Symposium, Technical Report FS-98-02, AAAI Press, Menlo Park, California*

Feldman, R., Gordon, I., Schneiderman, I., Weisman, O., & Zagoory-Sharon, O. (2010). Natural variations in maternal and paternal care are associated with systematic changes in oxytocin following parent–infant contact. *Psychoneuroendocrinology*, *35*(8), 1133-1141. https://doi.org/10.1016/j.psyneuen.2010.01.013

Floridi, L. (2017). Digital's Cleaving Power and Its Consequences. *Philosophy & Technology*, *30*(2), 123-129. https://doi.org/10.1007/s13347-017-0259-1

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, *30*(4), 681-694. https://doi.org/10.1007/s11023-020-09548-1

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 659-669. https://doi.org/10.1098/rstb.2006.2002

Gaudenzi (Forthcoming). Physical reality according to Bohr and intelligence according to Turing.

Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. *Journal of Artificial General Intelligence*, *5*(1), 1.

Guerrero-Rosado, O., & Verschure, P. (2021). Robot regulatory behavior based on fundamental homeostatic and allostatic principles. *Procedia Computer Science*, *190*, 292-300. https://doi.org/10.1016/j.procs.2021.06.039

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1-3), 335-346. https://doi.org/10.1016/0167-2789(90)90087-6

Harnad, S. (2018). Grounding symbolic capacities in robotic capacities. In *The Artificial Life Route to Artificial Intelligence* (pp. 187-214). Routledge

Hansell, M. (2000). *Bird nests and construction behaviour*. Cambridge University Press.

Haugeland, J. (1989). *Artificial intelligence: The very idea*. MIT press

Ingold, T. (1999). Three in one: on dissolving the distinctions between body, mind and culture. *French ("Trois en un: sur la dissolution des distinctions entre corps, esprit et culture") in Joulian F.(éd), Quand les chimpanzés ont dérobé la culture, CNRS editions*

Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. A. Sebeok (Ed.), Style in language (pp. 350-377). Cambridge, MA: MIT Press.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*(1-2), 1-19.

Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus artificial intelligence. *Frontiers in artificial intelligence*, *4*, 622364

Kupiec, J. (2019). Et si le vivant était anarchique: La génétique est-elle une gigantesque arnaque ?. Editions Les Liens Qui Libèrent.

Kurzweil, R. (2005). The singularity is near: When humans transcend biology. *Viking Penguin*.

Levesque, H., & Lakemeyer, G. (2008). Cognitive robotics. *Foundations of artificial intelligence*, *3*, 869-886

Liotti, G. (2005). *La dimensione interpersonale della coscienza*. Roma: Carocci.

MacLean, P. D. (1982). On the origin and progressive evolution of the triune brain. In *Primate brain evolution: Methods and concepts* (pp. 291-316). Springer US

MacLean, P. D. (1990). *The triune brain in evolution: Role in paleocerebral functions.* Springer Science & Business Media

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

Manning, A., & Dawkins, M. S. (2012). An introduction to animal behaviour. Cambridge University Press.

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological science*, *16*(1), 6-10.

Marković, D., Mizrahi, A., Querlioz, D., & Grollier, J. (2020). Physics for neuromorphic computing. *Nature Reviews Physics*, *2*(9), 499-510.

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). *The AI Index 2024 Annual Report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth Summer

Mollo, D. C. (2024). AI-as-exploration: Navigating intelligence space. arXiv preprint arXiv:2401.07964.

Murphy, R. R. (2019). *Introduction to AI robotics*. MIT press

Newen, A., De Bruin, L., & Gallagher, S. (Eds.). (2018). The Oxford handbook of 4E cognition. Oxford University Press.

Ogata, T., Takahashi, K., Yamada, T., Murata, S., & Sasaki, K. (2022). Machine learning for cognitive robotics

Panksepp, J. (1986). The anatomy of emotions. In Biological foundations of emotion (pp. 91-124). Academic Press.

Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press.

Panksepp, J. (2014). The core emotional systems of the mammalian brain: The fundamental substrates of human emotions. In *About a body* (pp. 14-32). Routledge.

Pavlov, I. P. (1941). Lectures on conditioned reflexes. Vol. II. Conditioned reflexes and psychiatry.

Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge University Press. https://doi.org/10.1017/ CBO9780511499968

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Schore, A. N. (2021). The interpersonal neurobiology of intersubjectivity. *Frontiers in psychology*, 1366.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, *3*(3), 417-424.

Seth, A. (2021). *Being you: A new science of consciousness*. Penguin.

Sieratzki, J. S., & Woll, B. (1996). Why do mothers cradle babies on their left? *The Lancet*, *347*(9017), 1746-1748. https://doi.org/10.1016/S0140-6736(96)90813-2

Skinner, B. F. (1938). The behavior of organisms. new york: Appleton-Century-Crofts. American Psychologist, 221, 233.

Steels, L. (2001). Language games for autonomous robots. IEEE Intelligent systems, 16(5), 16-22

Steels, L. (2015). *The Talking Heads experiment: Origins of words and meanings* (Vol. 1). Language Science Press.

Sutton, R. and Barto, A. (2014) Reinforcement Learning, MIT Press

Tegmark, M. (2019). Being human in the age of artificial intelligence.

Thompson, E. (2010). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press.

von Humboldt, W. (1999). *Humboldt:'On language': On the diversity of human language construction and its influence on the mental development of the human species*. Cambridge University Press.

von Uexküll, J. (1934). Streifzuge durch die Umwelten von Tieren und Menschen. *Ein Bilderbuch unsichtbarer Wellen*

Wiley, R. H. (1973). The strut display of male sage grouse: a" fixed" action pattern. *Behaviour*, *47*(1-2), 129-152.

Will generative AI transform robotics?. *Nat Mach Intell* **6**, 579 (2024). https://doi.org/10.1038/s42256-024-00862-2

Ziemke, T. (2022). Embodiment in cognitive science and robotics. In A. Cangelosi & M. Asada (Eds.), Cognitive Robotics (pp. 1-20). MIT Press.