Human and Non-Human Consciousness: Do They Share Common Characteristics?

Evangelos Koumparoudi

Sofia University, 115 Tsar Osvoboditel Blvd., Sofia, 1504, Bulgaria
vkoumparoudis@phls.uni-sofia.bg

Abstract. This study examines the possible common characteristics between human and non-human consciousness. It mainly addresses animal consciousness and, to a certain extent, intelligent AI. It provides an overview of the main theories regarding consciousness, more specifically those of neuroscience and cognitive science, and also their materialistic base at a neuroanatomical and neurophysiological level, emphasizing the role the prefrontal cortex plays, both in humans and animals. Then, it considers particular aspects of consciousness, such as emotion, and presents the three broad traditions considering human emotions, which are emotions as feelings, evaluations, and judgments, as well as studies on animal emotions. Then, it continues with the proposed models of metacognition and memory to deepen the analysis regarding common characteristics of human and non-human consciousness. It also touches on the platform theory, which may bridge human, animal, and AI consciousness, although this theory is under consideration. It ends with references to animals’ social behavior, their interactions with humans, their possible ontogenic proximity as expressed in biolinguistics, and the findings of computational ethology, which help to establish models of mental human disorders. The study concludes that findings support proximities between humans and animals, consciousness at the level of neurophysiology, and emotion and metacognition. Contrary to animals and AI, human consciousness is more complicated and far from cybernetic and computational models since it is linked with various kinds of malleability, reconsolidation, neural plasticity, different conceptions of emotions, and certain mental pathologies.

Keywords: neuroscience, emotion, metacognition, memory, platform theory, biolinguistics, computational ethology

Article history:
The article was submitted on 29.08.2023
The article was accepted on 20.09.2023


© Koumparoudis E., 2023
This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode
Человеческое и нечеловеческое сознание: имеют ли они общие характеристики?

Евангелос Кумпарудис

Софийский университет,
Болгария, София, 1504, б-р Царя Освободителя, д. 115
vkoumparoudis@phls.uni-sofia.bg

Аннотация. В данном исследовании рассматриваются возможные общие характеристики человеческого и нечеловеческого сознания. В основном статья направлена на анализ сознания животных и, в определенной степени, разумного искусственного интеллекта (ИИ). Представлен обзор основных теорий, касающихся сознания, в частности теорий нейробиологии и когнитивной науки, а также их материалистической основы на нейроанатомическом и нейрофизиологическом уровне, подчеркивается роль, которую играет префронтальная кора как у людей, так и у животных. Затем рассматриваются отдельные аспекты сознания, такие, как эмоции, и представлены три широкие традиции, анализирующие человеческие эмоции, а именно такие, как чувства, оценки и суждения, а также исследующие эмоции животных. Затем предложенные модели метапознания и памяти используются для углубления анализа общих характеристик человеческого и нечеловеческого сознания. Это также затрагивает теорию платформ, которая имеет возможность соединить сознание человека, животного и искусственного интеллекта. Хотя эта теория находится еще в стадии обсуждения. Исследование заканчивается ссылками на социальное поведение животных, их взаимодействие с людьми, их возможную онтогенетическую близость, выраженную в биолингвистике, а также на открытия компьютерной этологии, которые помогают создавать модели психических расстройств человека. Результаты исследования подтверждают близость между людьми и животными, их сознания на уровне нейрофизиологии, эмоций и метапознания. Однако, в отличие от животных и искусственного интеллекта, человеческое сознание более сложное и далекое от кибернетических и вычислительных моделей, поскольку оно связано с различными видами податливости, реконсолидации, нейронной пластичности, с различными концепциями эмоций и некоторыми психическими патологиями.

Ключевые слова: нейронаука, эмоции, метапознание, память, теория платформ, биолингвистика, вычислительная этология

История статьи:
Статья поступила 29.08.2023
Статья принята к публикации 20.09.2023


Introduction

In this paper, my main goal is to present the possible common characteristics of human and non-human consciousness, mainly focused on animal consciousness. However, some brief mention of AI will be incorporated as parallel to the structure. I will start with the various types of consciousness as proposed by philosophers and
cognitive scientists to better conceive what consciousness, in general, is and what differentiates human and animal consciousness. Then, I will consider the neuroanatomy and neurophysiology of consciousness to establish the materialistic basis of the phenomenon. After that, I will address in greater detail certain aspects such as emotion, metacognition, and memory. I will thereby address the question as to whether animals share common characteristics with humans as regards cognition and if this could be a basis for the understanding of their particular traits.

Additionally, I will present a theory under consideration, the so-called digital platform theory, which tries to bridge human, animal, and AI consciousnesses. Finally, I will proceed to a more concrete basis concerning recent findings in animal social behavior and interaction with humans, their possible ontogenic proximity and the impact of those findings on biolinguistics and the theory of evolution. Lastly, I will refer to computational ethology and how, by examining complex, high-dimensional behaviors generated in animals, we can construct models of humans, mainly of their psychiatric disorders.

Types of Consciousness

Consciousness has had diverse meanings through the years. Aristotle, for example, believed that the center of the soul is the heart; therefore, the center of human reason of the *zoon logon exon* (animal with Reason) could be found somewhere in the chest. Many years later, Rene Descartes placed the soul in the pineal gland; therefore, the center of human reason was now our brain. This assumption created the mind-body problem, Cartesian interactionism, the interplay between our *res extensa* and *res existensa*. These today may seem antiquated, although they are essential in the establishment of the Western philosophical canon. Consciousness ordinarily means not being asleep or in a coma and, more profoundly, perceiving the characteristics of an environment [1; 2]. Ned Block, in 1995 [3], introduced the notion of access consciousness. This is connected with the capacity to action or speech as interpretations of mental representations. Animals, according to this point of view, may have consciousness since language is not a prerequisite, but from another point of view, this capacity of animals is different from function consciousness and from a qualitative and phenomenal feeling of what situations “are like,” which Thomas Nagel and David Chalmers describe in their *What Is It Like to be a Bat* [4] and *The Conscious Mind: In Search of a Fundamental Theory* [5].

Phenomenal consciousness is deeply connected with the subjective experience a human may have. It remains a question of whether an animal has phenomenal consciousness, and, as regards AI, this is one of the mysteries in the era of intelligent AI. AI was started as a project in 1956 by Marvin Minsky, Claude Shannon, Ray Solomonoff, Nathaniel Rochester, John McCarthy, and others, who were the pioneers in the establishment of proto-AI theory [6]. Before them, Alan Turing introduced the terms binded and double-binded Turing machine to describe the first forms of computers and, of course, the Turing Test to differentiate human and non-
human (computer) bits of intelligence. Many years later, Searle proposed the Chinese Room Test and made some propositions about AI as to whether it is superior or inferior to a computer. In brief, the criterion of superiority is the phenomenal consciousness. It refers to the qualitative, subjective, or phenomenological aspects of conscious experience, sometimes identified with qualia (experiences such as colors, tastes, noises, and other sensations) with their distinctive character [7]. Another main philosophical and technological achievement was the introduction of the notion of computational neuroscience in the late 1980s by Patricia Churchland, Christof Koch, and Terrence J. Sejnowski [8] and the creation of the first neuronal networks in computer science, today known as RNR (Recurrent Neuronal Networks), which are used in machine and deep learning processes. Today, though we may face some severe ethical and ontological dilemmas regarding AI, it should not scare us, for these technologies, even if they may have a very advanced form and capacities, have been developed over time. Cofoey and Smith [9] insist on this subjective experience, which only humans have, but unquestionably not animals or AI in a broader view.

As regards animal consciousness, what is it? Are animals conscious, and, as a result, should particular ethical theories and legislation be implemented, as well as strict rules regarding their well-being and welfare? The term sentience comes from the Latin senti-em, which is to feel, to be conscious of something. According to many authors, animals feel pain, suffering, and enjoyment [10—12]. In 2012, a Declaration of Consciousness was made in Cambridge, according to which all animals — mammals, birds, fish, octopuses, and cephalopods — have consciousness. At the European level, there were the first discussions for a new ethical and legal framework in 2014 in France, 2019 in the EU, and 2022 in the UK. Today, all of the EU state members, as well as the UK, define animals as living beings with sentience. We now proceed to two other theories of consciousness: higher-order thought theory and self-consciousness. According to the first, animals do not have consciousness since a conscious state is a state whose subject is, in some way, aware of being in it [13]. The second, as introduced by DeGrazia [11], distinguishes three forms of self-awareness: bodily, introspective, and social, and he argues that some animals may be self-aware in the bodily sense but not introspectively or socially; subsequently, they may have consciousness.

**Neuroanatomy and Neurophysiology**

Consciousness has a material base beyond its anthropological and historical context. For example, the understanding of everyday practices and cooperation between the members of a team or the understanding of history are indeed not strictly limited to the brain. On the other hand, many neuroanatomical and neurophysiological models have been proposed to grasp the phenomenon of consciousness; I will briefly refer to some of them. According to Endelman and Tononi [14], the parts of the brain that play an essential role in consciousness are the thalamocortical system and, more precisely, the prefrontal cingulate and parietal
cortices. Later, Nir and Tononi, Tononi, and Koch [15; 16] emphasized the function of connectivity in the cortical system and established a theory of information integration. A model that is close to theirs is that of Dehaene and Changeux [17], who propose that the phenomenon of consciousness is based on a global neuronal workspace, which is a group of cortical pyramidal neurons with an excitatory function with long-range cross-cortical axons. However, are there homologies between humans and animals? Researchers such as Baars [18] insist that the thalamocortical system is central to mammals, and similar functions can be found in EEG, NREM, and REM sleep.

**Emotion**

The definition of human emotions is epistemologically complicated. In order to proceed to a definition and modeling from the spectrum of philosophy of medicine and biology, we should follow a prescriptive definition far from generalizations. Three broad traditions try to define human emotions: those of evaluation, motivation, and feeling [19]. In order to clarify better the three traditions, I will analyze each category separately; before that, I will refer to two historical models of human emotions as possibly based on human-animal relations. The first is that of Charles Darwin, in his *The Expression of the Emotions in Man and Animals* [20], who gives us an evolutionist account of emotions. For him, human emotions are homologs to those of animals, and the basic emotions are anger, fear, surprise, and sadness, which are common among species and cultures. The second is the James-Lange theory, introduced around the late 1890s, which conceives emotions as feelings. According to it, emotions are feelings constituted by perceiving physiological changes; they are not eternal sacred psychic entities. These changes in physiological conditions are linked with autonomic and motor functions [21. P. 449; 22. P. 189–190]. The proposed model is summarized as **Stimulus → Autonomic Arousal → Conscious Feeling.** Cannon and Bard, in the 1920s [23], established a theory of emotions as feelings different than that of James and Lange. For them, if emotions are the perception of bodily changes, entirely dependent on intact sensory and motor cortices, they also revealed that removing the cortex does not eliminate emotions. They also explored the effects of brain lesions on the emotional behavior of cats, more precisely the sham rage, which is linked with sudden inappropriate, ill-directed anger attacks. Animal emotions are human homologs; the introductory part of the brain responsible for emotions is the hypothalamus, and the model consists of **Stimulus → Hypothalamus → Conscious Feeling + Autonomic Arousal.**

In the 1960s, CD Broad, Errol Bedford, Antony Kenny, and, more recently, Robert Salomon, Jerome Neu, and Martha Nussbaum proposed that emotions are evaluations or judgments. According to Kenny [24], “If emotions have intentionality … there are internal standards of appropriateness of an emotion … just in case its formal object is instantiated … But feelings are not the kinds of things that can enter conceptual relations with formal objects. So, to be properly
embedded in conceptual relations, emotions need to be or involve cognitive evaluation. Judgementalism is based on the model of: Is that an emotion \( E \) is a judgment that the formal object of \( E \) is instantiated (by some particular object \( X \)). Deigh [25] attacks judgmentalists, accusing them of a lack of motivation; he insists on a phenomenological approach and explores the emotions of animals and infants. Solomon [26. P. 105—106] responds to him by claiming that emotions have a core desire that makes them motivational (e.g., fear encloses the core desire to flee). Finally, Nussbaum [27. P. 45] supports animals and infants being phenomenologically salient since they involve pre-linguistic and non-linguistic acceptance of how the world seems. Additionally, some theories assert that emotions are motivations connected with appraisal and affective science; in short, they describe how significant a situation is for an individual. Arnold, in the 1960s, spoke about eliciting circumstances, good/bad, present/absent, and easy to attain/avoid [28]. Lazarus 1991 writes about six possible categories of appraisal: a) goal relevance, b) goal concurrence, c) type of ego-involvement, d) blame or credit, e) coping potential, and f) future potential. Scherer et al. [29] distinguished between sixteen dimensions of appraisal, labeled stimulus evaluation checks (SECs), which can be grouped into four classes: appraisals of relevance, appraisals of consequences, appraisals of coping potential, and appraisals of normative significance. The model, in short, is Stimulus → Autonomic Arousal → Appraisal → Conscious Feeling. Motivational theories are phenomenological and non-phenomenological. For Deonna and Teroni [30], emotions are feelings of action-readiness; an experience of the dog is dangerous insofar as it is “an experience of one’s body being prepared” for avoidance. Contrary to them, Scarantino [31] argues that emotions are causes of states of action readiness, which may or may not be felt, and a general direction for behavior by selectively potentiating coherent behavioral options.

Two more proposed models of human emotions account for the intermediation of the whole body or at least try to reject Cartesian interactionism and the mind-body split or presuppose a high level of brain-body cooperation. The first is a single-system model, such as that of Cannon and Bard, proposed by Antonio Damasio, who criticized Rene Descartes’ famous cogito ergo Sum. For Damasio, we are more emotional beings with reason than reasonable emotions. After the conduction of various fMRIs, he proposed a model that places the prefrontal cortex as central to emotion regulation. There is a somatic marker hypothesis in which there are physiological reactions like shifts in the autonomic nervous system activity that tag previous emotionally significant events. These feelings are important in decision-making.

Furthermore, there is an as-if loop through which the brain areas that evaluate a stimulus (the amygdala and the prefrontal cortices) can directly signal the somatosensory cortices instead of triggering bodily activity [32]. Contrary to Damasio, Terry Davinson et al. [33] propose a dual-system approach, the so-called valence asymmetry model. The prefrontal cortex has two systems, one of approach
and one of withdrawal. The first is responsible for facilitating appetite behavior and generating particular types of positive affect that are approach-related, emotion moving towards a desired goal. In contrast, the second facilitates the withdrawal of an organism from sources of aversive stimulation and/or organizes appropriate responses to threat cues. This system also generates withdrawal-related negative emotions such as disgust and fear.

To conclude, there is no universal definition but three broad traditions: emotions as feelings, as evaluations/judgments, and as motivation. There are conflicting theories, such as Cannon-Bard, James-Lange, Deigh, and those of judgmentalists. Finally, concerning modeling, there are single-system and dual-system, e.g., Cannon-Bard and Damasio and Davinson.

Experiments have been performed in sheep and rats that reveal a specific background for animal emotions. Sheep not only have emotion-related responses, which could be considered depending on stimulus-response processes, but they also experience complex emotional states. Furthermore, they may experience a wide range of emotions, including fear, rage, despair, and boredom, via their sensitivity to suddenness, unfamiliarity, unpredictability, and discrepancy from expectations, controllability, and social norms [34]. Finally, rats suggest that the anticipatory behavior of an animal to a positive event is linked with positive emotional states [35].

**Metacognition, Memory, and Platform Theory**

One of the other equally important issues of animal consciousness is whether they have metacognition. Experiments have been performed in dolphins and monkeys whose evidence is supportive, but other experiments show various criticisms of animal metacognition. For dolphins, we classify low and high-frequency tones — similar to humans through a paddle [36], and for monkeys, there is a classification of visual stimulus and memorization [37]. Criticism comes from the hide-and-seek play of baboons, as the repeat key used in the experiment did not improve their performance. Consequently, they could not use any extra information in performing their task. This capacity has not been confirmed in other animal species. [38; 39]. Additionally, Hampton [40] proposes that during animal training, we do not necessarily have high-level metacognitive processes.

Concerning how human memory works, we refer to Hebb and his idea of “neuronal assembly,” in which the trace is presented as the increase of connective strength between populations of interconnected neurons and parallel cell co-activation. According to the positivist accounts of the 1940s, the term *memory* was replaced by the term *learning*; the cybernetic theory of that age appealed more to the way information is transmitted between humans (and animals). Atkinson and Shiffrin proposed the “modal model” in which we find terms such as sensory memory (SM), short-term memory (STM), and long-term memory (LTM); the last, seen as autobiographical memory, is closer to the “trace.” In 1997, Nadel and Moscovitch gave us the “multiple trace theory”; they highlight that the episodic character of
memory trace is placed in the linked ensembles of the hippocampal complex. The existence of the trace depends on semantic assemblies at the time of trace reactivation. In recent years, we have the emergence and broadening of this model in the so-called “reconsolidation hypothesis,” which suggests the malleability of the trace. It emphasizes the plasticity of the human brain and, thus, the approach of the trace as not overdetermined by its inscription, giving rise to a subject who acts in an autopoietic way by creating new traces from the inaugural traces. Finally, as for phenomenological approaches, the debate within this perspective revolves around Merleau-Ponty and his considerations on the subject, who always participates in the act of perceiving. The perceptual object is a product of the encounter with the world. This led many scientists to explore brain function during information assimilation experimentally. They found that the energy consumption of the brain is only 5% when perceiving. The other 95% is linked with the so-called “default mode of the brain,” or the amount of the information that the retina can handle, and, therefore, the brain ascribes a certain coherence to the information we perceive so this coherence can be linked with perceptual traces. In phenomenology, there are the “bottom-up” and “top-down” theories. In the first, the trace is seen as an impression that remains in the mind; in the second, previously inscribed traces shape our perception [41].

Animals, for some researchers, do not have episodic memory. Therefore, they are incapable of mental time travel; they are stuck in the present [42]. According to others, various breeding and mating behaviors may reveal a time travel [43; 44]. The debate concerns whether they can make future projections and whether future projections exist in infants and other non-verbal humans [45]. The latter is the central debate of the platform theory that we will briefly address. The platform theory, although it is new and under much consideration, may propose a bridge between human and non-human consciousness, even AI.

In the platform theory proposed by Zlomuzica and Dere [46], the phenomenon of consciousness is connected with an effortful action; therefore, we do not have consciousness when we, for example, drive a car. The model requires a confident presence of alert and responsive waking; conscious cognitive operations require the preservation of mental representations, so this theory is close to the representational theory of mind. The consciousness starts and ends with this kind of conscious cognitive operation. Animals do not have that capacity and are incapable of future time travel. Experiences, sensations, and mental time travel may be the content of consciousness but are not the consciousness itself. Rather than searching for human-like consciousness in animals, they suggest a series of behavioral tasks (the main experiments were done with rats in mazes of a certain complexity) to demonstrate conscious cognitive operations in animals. The definition of consciousness should apply to humans, animals, and artificial intelligence.

Social Behavior, Interaction, and Computational Ethology

Social behavior in animals is expressed in the formation of teams. Here, it is important to distinguish between automation and integration. Simply, it is different
for a part of an animal team to be a simple automaton following the other team members than to be conscious and cooperate in an integrated form. Additionally, it is important to have individual recognition and not solely be a part of a total [47; 48]. As for their interactions with humans, we tend to attribute human-like emotions and characteristics to our pets, especially dogs. For example, my dog is sad, or my dog is clever like a human, etc. [49]. Another model is that of symbiotic exchanges, for example, infant and mother; this exchange may have a positive or negative impact on the relationship [50]. Nagasawa et al. (2015) [51] revealed that the hormone oxytocin is produced when we have an intimate relationship with a dog. Finally, there may be an ontogenic proximity between humans, animals, and birds. The FOXP2 gene is common between species and is responsible for verbal capacity. According to the domain of biolinguistics, this verbal capacity may be common so that we could shape an integration hypothesis of a common ontogenic verbal proximity. Today, we know that the gene mutation in humans is connected with developmental verbal dyspraxia and childhood apraxia of speech [52; 53]. A different pathway is continuity and discontinuity theories; during evolution, humans co-evolved with animals, or they followed a different evolution separately from animals.

Recent findings in the domain of computational ethology try to establish bridges between human and animal behavior. Experiments use AI technologies, advanced video recording and monitoring, and VR technologies. They deal with how complex high-dimensional behaviors are generated by neural systems, temporal dynamics, and neural activity. They then proceed to quantifying free naturalistic movements of decision-making in complex naturalistic scenarios [54]. This modeling leads to possible explanations of human, mainly mental, disorders. For example, thigmotaxis in animals is a defensive animal measure of anxiety where they stay close to walls rather than maneuver in open spaces, which is possibly connected with social phobia and autism in humans [55]. Furthermore, transdiagnostic models of psychiatric disorders, which deal with dysfunction in learning and decision-making, have also been explored [56]. Finally, research domain criteria lead to frameworks of mental disease classification [57] and the development of new drugs [58].

Conclusions

Human consciousness has been analyzed thoroughly through the years concerning its anthropological and philosophical dimensions and neuroscientific and developmental aspects. In animals, things follow a possible parallel way; at a neuroanatomical level, the thalamocortical system plays a central role in humans and animals. We refer to mammals and those animals with a fully developed prefrontal cortex. However, in humans, the suggested models require a more complex modeling, that of information integration. As for AI, the fundamental dilemma is whether it is capable of a kind of consciousness far from computational modeling and requires the ability of “what is like” situations, namely phenomenal consciousness. The platform theory, briefly analyzed above, belongs to a representational theory of mind. The central argument is of working memory and the reaction of humans,
animals, and AI in alert states; it also highlights animals’ incapability of future time travel. When we drive or perform different tasks we may as humans be unconscious, but, as we elaborated, human memory is very complex and cannot be reduced to simple cybernetic or computational models since the notion of the trace is connected with various kinds of malleability, reconsolidation, and neural plasticity, as well as some instances of mental pathologies. The latter is also connected with the philosophy and biology of emotions. Human emotions have at least three discrete traditions: they can be conceived as feelings, evaluations or judgments, and motivations. Animals may also have feelings, for example, sheep and rats. We could say that they can fit the conception of emotions as feelings since they feel pain, enjoyment, etc.; this is also the reason for defining animals as sentient beings and for the changes in well-being and welfare legislation, protocols, and regulations. As for the second category, evaluations, and judgments, their capability of metacognition is under much consideration; some findings may support animal metacognition and some others that do not verify such an assertion. As for the last, motivation, their social behavior and interaction with other members of an animal team or humans may be a basis for creating some further assumptions, but we still may be at a very preliminary level of claiming that animals have emotions which may be motivations, at least as occur in humans.

Furthermore, from the findings of biolinguistics and computational ethology, they preserve a common place of understanding rather than distancing our interrelation. In any case, we are different species that may share common characteristics. This also goes for AI, but a harmonious co-existence presupposes an asymmetrical mutual understanding. The relation could be Other-directed. I may introduce an ethical dimension here, concerning Emmanuel Levinas, but even for the philosophy of biology and medicine, we could see ourselves not as the dominant We, I, etc., but the animals and AI as the Other, permanently withdrawing from the third, and placing animal and AI ethics as fundamental in our relationship, far from anthropomorphic and zoomorphic fallacies.

References


[35] Harst JVE. Standard housed rats are more sensitive to rewards than enriched housed rats as reflected by their anticipatory behavior. Behavioral Brain Research. 2003;142(1—2):151—156.


**About the author:**

Koumparoudis Evangelos — PhD in Philosophy, Post-Doctoral Researcher in Philosophy of Medicine and Biology, Faculty of Philosophy, Sofia University, Sofia, Bulgaria (e-mail: vkoumparoudis@phls.uni-sofia.bg). ORCID: 0000-0002-1068-4376

**Сведения об авторе:**

Кумпарудис Евангелос — кандидат философских наук, пост-докторант в области философии медицины и биологии, философский факультет, Софийский университет, София, Болгария (e-mail: vkoumparoudis@phls.uni-sofia.bg). ORCID: 0000-0002-1068-4376