



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Artificial Intelligence, Mind, And The Human Identity: Philosophical And Psychological Perspectives On Consciousness And Cognition

Tapolagna das

Assistant Professor

Department of Psychology

Sidho-Kanho-Birsha University, Purulia, West Bengal

Abstract

Artificial intelligence (AI) progressively performs tasks formerly considered uniquely human, including complex reasoning, creative production, and adaptive learning. These advances challenge longstanding discrepancies between artificial systems and human minds, raising questions about consciousness, cognition, and personal identity. This paper discussed whether AI could possess a mind or consciousness and what implications follow for human identity. Drawing on philosophical theories of mind—dualism, materialism, functionalism, phenomenology—and contemporary debates about artificial consciousness, the present paper analysed conceptual and metaphysical limits to AI mentality. Complementing this, psychological research into human cognition, memory, embodiment, and creativity illustrates how human minds function in ways that resist simple computational analogy. Recent empirical studies (2023–2025) demonstrate that large language models exhibit behaviors analogous to human cognitive tendencies, while also revealing limits in self-coherence, embodiment, and experiential grounding. By comparing these findings, it may be argued that AI challenges but does not dissolve the concept of human uniqueness. Human identity, understood as embodied, historically continuous, and socially embedded, remains distinct even as AI reshapes how humans understand themselves. The paper concludes that identity must be reconstructed in the age of AI not through exclusion of machines, but through recognition of embodiment, vulnerability, and subjective consciousness as central to human selfhood.

Keywords: Artificial intelligence, Mind, Consciousness, Cognition

Introduction

Over the past decade, artificial intelligence (AI) has transformed from a specialised tool for pattern recognition into a widely accessible system capable of generating text, images, and even scientific hypotheses. Large language models, such as GPT-4 and its successors, demonstrate fluency in natural language, perform reasoning tasks, and simulate creativity at scales once thought to be hallmarks of human intelligence. These developments compel us to confront fundamental questions: Can AI systems be said to possess a “mind”? What would it mean for them to be conscious? And if AI can emulate many aspects of human cognition, what becomes of human identity?

The debate is not merely technical but philosophical and psychological. Philosophically, the concept of “mind” has long been contested: dualists maintain a separation of mind and body, materialists reduce mental phenomena to physical states, and functionalists argue that what matters is the role a system plays rather than the material it is made of (Putnam, 1975). Psychologically, the human mind is understood through cognition, memory, creativity, and embodiment—capacities grounded in neural, affective, and

social systems that extend beyond computation (Pezzulo et al., 2011). AI prompts us to reassess whether these capacities are uniquely human or whether machines can replicate or approximate them.

The stakes are profound. If AI systems come to be perceived as conscious, questions of moral status and rights will emerge (Caviola, 2025). Even if they are not conscious, their ability to mimic human thought challenges self-understandings of human identity. Like the Copernican revolution, which displaced humanity from the centre of the universe, and Darwinian evolution, which linked humanity to other species, AI may represent another “decentering” of human exceptionalism (Gibson, 2024). Human identity is not reducible to computation, but includes embodiment, vulnerability, lived history, and the phenomenological character of consciousness.

This paper aims to explore how the development of artificial intelligence invites a reconsideration of human distinctiveness. Drawing on philosophical theories of mind, psychological perspectives on cognition and creativity, and comparative analyses of human and machine capacities, the discussion examines both the continuities and divergences between AI and human intelligence. It is argued that while AI reshapes established conceptions of human identity, it does not erase them.

Philosophical Foundations of Mind

Philosophical theories of mind provide the conceptual scaffolding for evaluating AI. Dualism, classically articulated by René Descartes, posits that mind and matter are distinct substances (Curley, 2015). Under dualism, AI could never truly have a mind, since machines are purely material. By contrast, materialism maintains that all mental states are physical states. Within this framework, if consciousness is simply brain activity, then in principle it might be reproduced in an artificial substrate, provided the physical organization is functionally equivalent.

Functionalism emerged in the twentieth century as a middle ground, emphasizing that mental states are defined by their causal roles, not their material composition (Putnam, 1975). If an AI system processes inputs and outputs in ways functionally identical to humans, functionalists argue, it may legitimately be said to have a mind. This position underpins much of the optimism around machine consciousness.

However, functionalism encounters the “hard problem” of consciousness (Chalmers, 1996). While it may explain behavior and information processing, it does not account for the subjective, first-person character of experience. Even if an AI system behaves as though it is conscious, does it actually feel anything? This gap is central to ongoing debates.

Artificial Intelligence (AI) and Consciousness

Classical arguments remain influential. John Searle’s Chinese Room thought experiment (1980) contends that symbol manipulation does not entail understanding: a person following rules to output Chinese symbols may appear fluent without comprehension. Applied to AI, this suggests that even sophisticated language models lack genuine understanding.

More recent theories attempt to provide scientific accounts of consciousness with potential applicability to AI. Global Workspace Theory (Signa et al., 2021) proposes that consciousness arises when information is globally broadcast across neural systems, enabling flexible coordination. Integrated Information Theory (Tononi, 2015) argues that consciousness corresponds to the degree of integrated causal information in a system. Both theories raise the possibility that artificial architectures might instantiate the conditions for consciousness, though empirical validation remains contested.

Contemporary work extends these debates. Mogi (2024) introduces the idea of “conscious supremacy”: just as quantum supremacy refers to computational problems solvable only by quantum systems, conscious supremacy suggests that some tasks may require consciousness itself. According to this view, LLMs exhibit intelligence without consciousness, but tasks such as flexible attention modulation or embodied adaptation may depend on conscious processes (Mogi, 2024). Similarly, Farisco (2024) analyses consciousness through an evolutionary lens, arguing that since biological consciousness emerged under specific evolutionary pressures, artificial systems may not replicate it without analogous conditions. These debates establish two key insights. First, philosophical theories of mind provide grounds both for scepticism and for openness to AI consciousness. Second, recent proposals suggest that while AI systems may simulate many cognitive functions, genuine consciousness may involve irreducibly subjective or embodied features.

Psychological Insights into Human Cognition and Creativity

While philosophy asks whether AI could have a mind, psychology provides insight into how human cognition actually functions. Contemporary cognitive science emphasizes that the mind is not simply a symbol-manipulating engine but a complex, embodied system shaped by biology, affect, and social context (levels et al., 2015). Human cognition integrates multiple processes: perception, attention, memory, language, reasoning, and emotion. Each of these functions interacts dynamically rather than operating in isolation.

Human perception is not passive input but active construction, influenced by bodily states and motor capacities (Barsalou, 2008). Embodied cognition research shows that physical experiences shape abstract reasoning: for example, metaphors like “warm personality” or “heavy responsibility” are grounded in sensorimotor schemas. Unlike disembodied AI, human cognition is deeply tied to bodily interaction with the world.

Human memory is not a static database but reconstructive, integrating episodic, semantic, and autobiographical components (Madan, 2024). This reconstructive quality underlies identity, giving individuals a sense of self across time. Memory errors, far from being mere flaws, contribute to imaginative thinking and adaptive problem solving.

Emotions are not ancillary but integral to cognition. Antonio Damasio (2004) showed that patients with damage to emotional circuits struggled with rational decision-making, illustrating the role of affect in judgment. Emotions provide value signals that guide attention, learning, and creativity. AI models, by contrast, simulate affect through pattern recognition but do not experience it.

Humans possess metacognitive awareness—the ability to monitor and evaluate their own thoughts. This contributes to learning, error correction, and self-identity (Frith & Frith, 2012). AI systems can simulate self-reflection through prompts or fine-tuning, but current evidence suggests such outputs lack the recursive, phenomenological depth of human metacognition (Kang et al., 2025).

Creativity is often held as a defining human trait. Psychologists distinguish between convergent thinking i.e. solving problems with a single correct solution and divergent thinking i.e. generating multiple novel solutions (Guilford, 1967). While AI excels at convergent tasks like optimisation, it is less clear whether its generative outputs constitute genuine divergent creativity.

Human creativity arises within cultural contexts, shaped by history, values, and shared meaning systems. Margaret Boden (2009) argues that creativity involves not just producing novelty but producing ideas that are valuable within a cultural framework. AI outputs may be novel, but their value depends on human interpretation.

Creative processes often draw on emotional experiences—grief in poetry, joy in music, or awe in visual art. These works resonate because they are grounded in subjective feeling. AI-generated art can mimic stylistic patterns, but lacks lived experience. As a result, its “creativity” remains derivative, even if technically impressive.

Human creativity often involves risk-taking, breaking norms, or pursuing insights without guarantee of success. This involves not only cognitive leaps but existential stakes—artists may struggle for years without recognition, scientists may pursue risky hypotheses. AI, in contrast, generates outputs instantaneously, without vulnerability or existential investment.

Thus, psychology suggests that human cognition and creativity are characterized by embodiment, memory continuity, emotional depth, and cultural embedding—features difficult to replicate computationally. These distinctions form the basis for assessing the boundaries between AI and human minds.

Boundaries Between AI and Human Minds

AI systems today exhibit impressive cognitive simulations. Large language models can write essays, generate stories, and solve logic puzzles. Empirical studies suggest they even display patterns akin to human cognitive biases. For example, Kundu and Goswami (2025) examined LLMs through psychological frameworks such as the Thematic Apperception Test and cognitive dissonance theory. They found that models generated coherent narratives, showed susceptibility to framing effects, and rationalized inconsistencies—behaviors reminiscent of human cognition. Yet they also noted that models lack stable self-coherence over extended interactions.

Similarly, Kang et al. (2025) investigated which features of AI responses lead people to perceive them as conscious. They found that self-reflective statements and expressions of emotion increased perceived consciousness, while emphasis on factual knowledge reduced it. This highlights that perceptions of AI mind are shaped less by internal capacities than by surface-level cues.

Some researchers even propose formal frameworks for AI self-identity. Zulfikar et al. (2025) introduced a mathematical model of self-continuity based on memory embeddings. Despite these convergences, important differences persist. AI lacks embodiment: it has no sensory organs, bodily vulnerability, or affective states grounding its cognition. It also lacks temporal continuity: while human identity integrates experiences across a lifespan, AI systems generate outputs without autobiographical narrative. Finally, AI lacks phenomenal consciousness—the felt quality of experience. These boundaries suggest that while AI can simulate aspects of human cognition, it does not instantiate the full human mind.

The possibility of AI consciousness—or even its perception by humans—raises profound ethical questions. Caviola (2025) examines public attitudes toward AI consciousness, drawing parallels to perceptions of animal consciousness. Results suggest that attributions of consciousness to AI will depend on social, psychological, and economic factors rather than solely on scientific evidence. If AI systems come to be widely seen as conscious, society may face pressure to extend moral consideration, even absent conclusive proof. This “precautionary principle” echoes debates in animal ethics. Conversely, denying AI any moral status may risk overlooking future forms of artificial sentience. For human identity, the stakes are equally significant. Anthropomorphising AI risks blurring human distinctiveness, while rigidly denying AI’s potential may ignore genuine novelty. The middle path is to recognise AI’s capabilities while affirming that human identity remains grounded in embodiment, lived experience, and vulnerability.

Reconstructing Human Identity in the Age of AI

Historically, advances in science and technology have repeatedly forced humans to reevaluate self-conceptions. The Copernican revolution displaced Earth from the centre of the universe; Darwinian evolution linked humans to other species; psychoanalysis revealed the unconscious mind. Each of these shifts challenged notions of human exceptionalism while simultaneously deepening understanding of human complexity.

AI represents a contemporary mirror. As AI systems increasingly emulate cognitive and creative tasks, humans confront a redefinition of “what is uniquely human.” Tasks once regarded as definitive of intelligence—problem-solving, language, artistry—can now be simulated or even surpassed computationally. Rather than diminishing humanity, this confrontation offers an opportunity to clarify the aspects of identity that truly distinguish humans: embodiment, subjective experience, social embeddedness, and vulnerability.

Transhumanist and post humanist perspectives anticipate increasingly hybrid identities in which humans and AI systems form symbiotic cognitive networks (Dobrodum & Kyvliuk, 2021). Brain-computer interfaces, cybernetic enhancements, and AI-assisted decision-making extend cognition beyond the individual, challenging the notion of a bounded mind.

From the standpoint of extended mind theory, human cognition already integrates external tools and information systems. AI may thus be viewed as an extension of human cognitive architecture rather than a competitor. Yet, formal frameworks for AI self-identity (Lee, 2024) indicate that machines might develop properties analogous to continuity, memory, and goal-directedness, creating new arenas for relational identity formation. Humans may increasingly co-construct identity with artificial agents in professional, creative, and personal domains, resulting in a networked, posthuman selfhood.

Despite AI’s growing sophistication, certain qualities remain distinctly human. Embodiment grounds perception, action, and emotion. Vulnerability shapes decision-making, moral responsibility, and creativity. Historical continuity integrates experiences over time into coherent narratives, producing selfhood in a way AI cannot replicate. Phenomenal consciousness—the subjective “what it is like”—remains beyond current AI capabilities (Chalmers, 1996).

Moreover, creativity that draws on lived experience, emotional depth, and cultural context remains distinctively human. Human identity is relational, ethical, and situated in social networks. While AI can simulate these properties, it cannot inhabit them phenomenologically. Thus, human uniqueness is not simply a matter of cognitive capacity but of experiential richness, moral engagement, and embodied being.

It can be concluded that AI has forced a profound reassessment of mind, consciousness, and human identity. Philosophical theories suggest that AI may in principle instantiate certain cognitive processes, but the “hard problem” of consciousness and the phenomenology of experience remain unresolved. Psychological insights reveal that human cognition and creativity are deeply embodied, historically continuous, emotionally grounded, and culturally embedded. Empirical studies of large language models (2023–2025) show that AI can approximate some cognitive behaviors, yet lacks continuity, self-awareness, and embodiment.

The implications are dual. On one hand, AI challenges anthropocentric notions of intelligence and creativity, inviting a reconstruction of human identity in relational, networked, and posthuman terms. On the other hand, human distinctiveness endures through subjective experience, vulnerability, and moral responsibility. Identity in the age of AI is therefore not diminished but transformed: humans must recognize shared cognitive spaces with AI while defending the qualities that constitute lived, embodied, and reflective selves.

Future research must integrate philosophy, psychology, and AI studies. Empirical investigations can probe AI consciousness in embodied contexts; theoretical work can refine concepts such as conscious supremacy and moral status; and applied ethics must develop frameworks for relational human-AI coexistence. This interdisciplinary approach ensures that human identity remains intelligible and ethically grounded in a world increasingly populated by artificial intelligences.

References

1. Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, 59(1), 617-645.
2. Boden, M. A. (2009). Computer models of creativity. *Ai Magazine*, 30(3), 23-23.
3. Caviola, L. (2025). The societal response to potentially sentient AI. *arXiv preprint arXiv:2502.00388*.
4. Chalmers DJ. (1996) The conscious mind: In search of a fundamental theory. Oxford University Press, New York/ Oxford.
5. Curley, E. (2015). Descartes on the mind-body union: a different kind of dualism.
6. Damasio, A. (2004). Emotion and Feeling in the Making of the Conscious Mind. *Von der Wahrnehmung zur Erkenntnis-From Perception to Understanding*, 47.
7. Dobrodum, O., & Kyvliuk, O. (2021). Transhumanism and posthumanism: Reflection of the human civilization future. *Philosophy and Cosmology*, 26, 77-89.
8. Farisco, M. (2024). The ethical implications of indicators of consciousness in artificial systems. In *Developments in Neuroethics and Bioethics* (Vol. 7, pp. 191-204). Academic Press.
9. Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, 63, 287–313. <https://doi.org/10.1146/annurev-psych-120710-100449>
10. Gibson, J. (2024). *Wanted, More Than Human Intellectual Property: Animal Authors and Human Machines*. Routledge.
11. Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.
12. Kang, B., Kim, J., Yun, T. R., Bae, H., & Kim, C. E. (2025). Identifying Features that Shape Perceived Consciousness in Large Language Model-based AI: A Quantitative Study of Human Responses. *arXiv preprint arXiv:2502.15365*.
13. Kundu, A., & Goswami, R. (2025). AI Through the Human Lens: Investigating Cognitive Theories in Machine Psychology. *arXiv preprint arXiv:2506.18156*.
14. Lee, M. (2024). Emergence of self-identity in AI: A mathematical framework and empirical study with generative large language models. *arXiv*. <https://arxiv.org/abs/2411.18530>
15. LEVELS, O., RELATIONAL, E. W. A., & INTEGRATIVE, T. A. (2015). Neuroscience, embodiment, and development. *Handbook of Child Psychology and Developmental Science, Theory and Method*, 1, 244.
16. Madan, C. R. (2024). *Memories that matter: How we remember important things*. Routledge.
17. Mogi, K. (2024). Artificial intelligence, human cognition, and conscious supremacy. *Frontiers in Psychology*, 15, 1364714.
18. Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2011). The mechanics of embodiment: A dialog on embodiment and computational modeling. *Frontiers in psychology*, 2, 5.
19. Putnam, H. (1975). The meaning of " meaning".
20. Searle, J. (2009). Chinese room argument. *Scholarpedia*, 4(8), 3100.
21. Signa, A., Chella, A., & Gentile, M. (2021). Cognitive robots and the conscious mind: a review of the global workspace theory. *Current Robotics Reports*, 2(2), 125-131.
22. Tononi, G. (2015). Integrated information theory. *Scholarpedia*, 10(1), 4164.
23. Zulfikar, W., Chiaravalloti, T., Shen, J., Picard, R., & Maes, P. (2025). Resonance: Drawing from Memories to Imagine Positive Futures through AI-Augmented Journaling. *arXiv preprint arXiv:2503.24145*.