



Published: June 30, 2024

**Citation:** Dubey S, Ghosh R, et al., 2024. Redefining Cognitive Domains in the Era of ChatGPT: A Comprehensive Analysis of Artificial Intelligence's Influence and Future Implications, Medical Research Archives, [online] 12(6). <https://doi.org/10.18103/mra.v12i6.5383>

**Copyright:** © 2024 European Society of Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**DOI**

<https://doi.org/10.18103/mra.v12i6.5383>

**ISSN:** 2375-1924

## Redefining Cognitive Domains in the Era of ChatGPT: A Comprehensive Analysis of Artificial Intelligence's Influence and Future Implications

Souvik Dubey,<sup>1</sup> Ritwik Ghosh,<sup>2</sup> Mahua Jana Dubey,<sup>3</sup> Subhankar Chatterjee,<sup>4</sup> Shambaditya Das,<sup>1</sup> Julián Benito-León,<sup>\*5,6,7,8</sup>

<sup>1</sup> Department of Neuromedicine, Bangur Institute of Neurology (BIN) Kolkata, West Bengal, India

<sup>2</sup> Department of General Medicine, Burdwan Medical College, and Hospital, Burdwan, West Bengal, India

<sup>3</sup> Department of Psychiatry, Berhampore Mental Hospital, Berhampore, West Bengal, India

<sup>4</sup> Department of Endocrinology and Metabolism, Medical College Kolkata, Kolkata, West Bengal, India

<sup>5</sup> Department of Neurology, University Hospital "12 de Octubre", Madrid, Spain

<sup>6</sup> Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain

<sup>7</sup> Centro de Investigación Biomédica en Red Sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain

<sup>8</sup> Department of Medicine, Faculty of Medicine, Complutense University, Madrid, Spain

\*Corresponding author: [jbenitol67@gmail.com](mailto:jbenitol67@gmail.com)

### ABSTRACT:

**Background and Objectives:** Despite its extensive utilization, research on Chat Generative Pre-trained Transformer (ChatGPT)'s potential negative impact on specific cognitive processes is scarce. This article explores the widespread use of ChatGPT in educational, corporate, and various other sectors, focusing on its interaction with distinct cognitive domains such as attention, executive function, language, memory, visuospatial abilities, and social cognition.

**Methods:** A literature review was conducted using PubMed, identifying 256 articles, with 29 peer-reviewed articles analyzed after screening for relevance.

**Results:** The review emphasizes the extraordinary capabilities of the human brain, which often go unrecognized, and argues for the importance of maintaining and enhancing natural cognitive abilities using artificial intelligence tools like ChatGPT as an aid rather than a replacement. The findings highlight the advanced reasoning capabilities of ChatGPT, blending intuitive and deliberate cognitive processes.

**Conclusion:** Building a socio-cognitive architecture for collective human-machine intelligence has significant potential. While ChatGPT offers impressive capabilities, over-reliance on it for cognitive tasks can lead to the erosion of essential skills. It is crucial to find a balance between leveraging artificial intelligence's advantages and preserving our natural cognitive abilities, ensuring continuous practice and engagement in traditional cognitive exercises.

**Keywords:** ChatGPT; cognition; artificial intelligence.

## Background And Objectives

Human cognition functions are based on intricate, intriguing, interdependent, and crucially orchestrated crosstalk among subdomains of the cognitive sphere.<sup>1</sup> Copacetic connections among the subdomains, i.e., attention, executive function, language, memory, visuospatial, visuoperceptual, gnosis, and praxis, build human cognition's backbone.<sup>2</sup> Human beings act as monads of society, and the aspect of human cognition dispersed socially for integration of our social attributions has been termed social cognition. Social cognition encompasses self and non-self identification, emotion perception, empathy, and theory of mind.<sup>3,4</sup> Each component of social cognition entails a supreme excursion of achieved cognitive abilities in a coordinated fashion. Genetics, epigenetics, and environmental influences perpetually imbue social behavior modifications intending to balance ever-changing social norms, regulations, and customized social cognition for better and healthier social survival.

Human behavior, conditioned by an intricate interplay of personality, social cognition, learning, and past experiences, customarily steers in a patterned way determined by situational complexity. Hence, cognitive abilities, including social cognition and human behavior, have numerous determinants underneath. Each sub-domain evolutionary perspective and framework is further complex and undergoes multifactorial processing, which begins during infancy and then sustains several modifications throughout the lifetime by innumerable and variable sensory inputs, learning, and acquired experiences. Sensory inputs through various modalities, i.e., acquisition of information and processing of procured information, encounter thousands of modifications (i.e., trials, errors, and checkers' balance).<sup>5,6</sup> After achieving the custom-tailored information, it is further subjected to several cornerstones of cognitive abilities, i.e., learned experiences, situational analysis, concept building during the crosstalk, and the ultimate experience-dependant expression (i.e., verbal, gestural and/or by writing). The central executive network plays a role in the middle through established bidirectional connections with auditory inputs (via a phonological loop) and visuospatial inputs (via visuospatial sketchpad). This fact argues that constant multimodal sensory inputs with numerous bidirectional modifications are prerequisites for supreme executive functions. Human beings are privileged with the supreme cognitive ability of "thought," which requires the engagement of multiple connectomes in the brain at a point in time based on attention, knowledge/concepts (i.e.,

semantics), learning, experiences, memory, emotions, imageries, imaginations, and anticipations. "Thought" is the cognitive ability that requires the functioning of every subdomain at a precise point in time. We "execute" our thoughts through a "thought-action" framework, which demands an even more widely and complex network of planning, organization, and sequencing in addition to "thought" and motor networks. In this way, our cognitive circuits have become increasingly strengthened and fine-tuned with errors, error checks, and corrections. Writing skill is one of the most valuable, supreme, and finest modalities to express our "thoughts" in written form. It is presumably clear from the discussion above that writing skill is a direct reflector of composite cognitive functions of the human brain, and it requires constant and regular excursions of all cognitive domains only to sharpen cognitive circuits and brain function as a whole.<sup>7-9</sup>

Chat Generative Pre-trained Transformer (ChatGPT) is an artificial intelligence-powered language model developed by OpenAI and released in November 2022 that can generate detailed and extensive written information virtually on any topic within seconds. Scientists have already argued whether ChatGPT can be considered an author in scientific literature. Students, teachers, corporate houses, and people from various streams have started using ChatGPT to integrate information and writing. Discussions have already gained attention on ChatGPT, and some researchers also wanted to categorize this invention as a revolution. ChatGPT provides a plethora of integrated information on any given topic within seconds.

We aimed to elucidate the interaction between ChatGPT and specific cognitive domains, including attention, executive function, language, memory, visuospatial abilities, and social cognition, and to assess the potential future impacts of ChatGPT on these cognitive processes.<sup>10-12</sup>

## Methods

In crafting this exploratory review paper, we conducted a targeted literature review using the PubMed database with the following keywords (["Cognition" OR "Cognitive AND ChatGPT" OR "Chat Generative Pre-trained Transformer"]) up to May 20, 2024. We also hand-searched additional ChatGPT-specific articles using the reference list of the selected studies and relevant journal websites from 2022 to the current date for literature inclusion. To decrease publication bias, we invigilated the references of all studies potentially missed in the electrical search. Content experts also

searched the gray literature for any relevant articles.

The initial search yielded a total of 256 articles. After removing duplicates, we screened titles and abstracts for relevance, narrowing down to 29 articles for in-depth assessment. The selection process was carried out by two independent reviewers, with any discrepancies resolved through discussion or by consulting a third reviewer, ensuring a robust and unbiased selection.

This exploratory analysis enabled us to leverage a diverse array of sources to underpin our discussion and ensure a thorough examination of the presented topics.

## Results

Cultural modulation of neural mechanisms asserts that technological interactions significantly influence brain development and cognitive abilities. The use of artificial intelligence chatbots like ChatGPT for cognitive offloading may lead to underemployment of specific cognitive faculties, inhibiting their full maturation. This phenomenon is particularly relevant in the context of executive functions, where reliance on artificial intelligence for problem-solving can reduce cognitive effort and lead to long-term cognitive changes.<sup>13</sup>

Considering the cognitive adjustment hypotheses, three potential scenarios for the impact of pervasive ChatGPT use could be proposed: First, the cognitive zero-sum adjustment hypothesis suggests minimal interaction with neurocognitive abilities. Second, the cognitive-soft adjustment hypothesis posits that ChatGPT will be used as external resources for specific tasks, leading to potential underutilization of higher-order cognitive functions. Finally, the cognitive-hard adjustment hypothesis envisions a dystopian scenario where artificial intelligence chatbots completely replace higher-level executive functions, fundamentally altering human cognition.

There are numerous reports in the literature detailing how ChatGPT functions and what its prospects are. However, there is a notable lack of articles examining the potential adverse effects of ChatGPT on specific cognitive processes. Of the 29 articles that were ultimately selected, none systematically addressed the interaction between ChatGPT and distinct cognitive domains. Therefore, we will express our perspective on this subject, considering each primary cognitive domain, including attention, executive function, language, memory, visuospatial abilities, and social cognition.

**Attention domain:** Writing requires sustained attention, which improves with practice. In the era of ChatGPT, people have seemed to forget the degree of attention needed to construct a good write-up. The reinforcing circuit of attention involved in writing will be at stake. The primary acquisition and/or expression of any cognitive function demands attention, and anything that is expected to be related with potential to decay attention must not be welcomed, at least in every field.

**Executive function domain:** Writing is no longer considered solely a language function; it requires the engagement of executive networks. Executive functions, require the functioning of multimodal brain networks. By showing dependence on ChatGPT for every minor to significant write-up, there is a chance of losing one of the most critical executive functions (i.e., writing) soon. This executive loss will be something that will be worth seriously considering as it will be painful to see humans using ChatGPT even to resolve minor issues.

A study by Hagendorff et al.<sup>14</sup> investigated how large language models like ChatGPT exhibit human-like intuitive thinking and reasoning biases. The study found that:

- Intuitive Behavior: Larger language models displayed more human-like intuitive (system 1) thinking and cognitive errors, but this decreased significantly in ChatGPT models.
- Chain-of-thought reasoning: ChatGPT models effectively use chain-of-thought reasoning to improve accuracy, mimicking human system 2 thinking.

In cognitive reflection tests and semantic illusions, both ChatGPT-3.5 and 4 models showed higher accuracy and fewer intuitive errors compared to earlier large language models<sup>14</sup>. These findings highlight the advanced reasoning capabilities of ChatGPT, blending intuitive and deliberate cognitive processes.

**Language domain:** Writing is the most critical aspect of the language domain within the human cognitive sphere, contributing significantly to our dominance on Earth. Writing ability depends on basic comprehension, knowledge (semantics), and grammar, making it one of the most complex language subdomains. Furthermore, writing enhances other subdomains within the language domain. However, in the era of ChatGPT, where creating a manuscript is as simple as giving a command, the pursuit of excellent writing exercises and skills is diminished. This reduction in writing practice can negatively impact other cognitive

subdomains, further weakening the overall cognitive abilities.

**Memory domain:** Writing has an intriguing association with memory. In order to prepare a write-up, human beings must rely on semantics and episodic memories. Acquisition of semantics occurs through multimodal sensory inputs, and acquisition of episodic memory requires registration, encoding, rehearsals, consolidation, storage, and timely retrieval. Steps often have a bidirectional relationship with intentions to boost others and demand rehearsals and training, which means practice and exercise. Writing needs all the abovementioned components: emotion, imagination, and imagery. The authors want to express their concern that over-dependence on ChatGPT will silently snatch away the excursion of these abovementioned cognitive abilities from human beings.

**Visuospatial, visuperceptual, praxis, and gnosis functions:** As the authors described, writing extends beyond the boundaries of the language domain. It also involves a comprehensive network of visuospatial and perceptual abilities, imagery, and imagination. These elements encompass visuospatial attention, perception, and visual memory associated with thought. "Gnosis" refers to modality-specific recognition of "what," making it a critical component of writing. Praxis in writing is self-explanatory. Therefore, relying on a "ready-made" write-up by ChatGPT may undermine these supreme cognitive abilities, limiting room for deeper perceptions. The limitless and reckless use of ChatGPT, particularly among students, may hinder the development of these finer cognitive abilities associated with writing.

**Planning, organization, novelty, and creativity:** It is needless to mention that higher frontal lobar associates like planning, organization, sequencing, novelty seeking, and creativity are the delicate aspects that are not only essential for writing but also determine the quality of writing. The "made up" writing by ChatGPT, in all aspects, may blunt the finer tunings in the future, as these essential skills will no longer be required.

A recent study compared how humans and ChatGPT retell stories.<sup>15</sup> The study highlighted several key findings:<sup>15</sup>

- Creativity and Novelty in Retellings: Humans create 55-60% of novel words and concepts in each iteration of retelling. ChatGPT provides competent summaries with fewer changes in subsequent retellings.

- Word Count and Part-of-Speech Differences: Both ChatGPT and humans shorten stories, but ChatGPT's retellings are more consistent. Humans use more verbs, adverbs, and negations, while ChatGPT uses more nouns, adjectives, and prepositions.
- Concepts: Creativity, Stability, and Decay: ChatGPT uses fewer synsets initially but maintains higher synset density across retellings. Humans generate more novel synsets in later retellings.
- Age of Acquisition: Human retellings preserve words learned earlier in life, while ChatGPT increases the average age of acquisition in retellings.
- Affect Preservation: Both human and ChatGPT retellings maintain the emotional core of stories with high fidelity.

These findings reveal significant differences in language use and creativity between humans and ChatGPT in the context of retelling stories. Humans demonstrate greater variability and creativity, altering stories more significantly across iterations. They introduce new words and concepts, mainly focusing on verbs and adverbs, which reflect actions and emotions. Humans also use more negations, which suggests a nuanced understanding of context and narrative complexity.<sup>15</sup>

In contrast, ChatGPT's retellings are more consistent, maintaining a higher degree of stability in word choice and structure. ChatGPT favors nouns and adjectives, focusing on the entities and descriptive aspects of the narrative. The use of later-acquired vocabulary by ChatGPT indicates a different processing mechanism compared to humans, who rely more on early-acquired, more easily retrievable words.<sup>15</sup>

Despite these differences, both humans and ChatGPT effectively preserve the emotional core of the stories they retell. This emotional stability suggests that while the words and specific details may change, the underlying emotional narrative remains intact, providing a consistent affective experience for the audience.<sup>15</sup>

**The behavior domains:** Quest and thirst for knowledge, concepts, and wisdom have made us what we boast about ourselves today. The search for knowledge should be spontaneous, open, and not limited by boundaries and 'spoon-feeding.' In this way, brain networks are boosted, and synaptogenesis and neuroplasticity develop. Threading scattered and fragmented knowledge and concepts in various permutations and

combinations associated with novelty seeking towards more significant concepts and inventions is the supreme-most cognitive ability considered a monopoly of human beings on Earth. Most importantly, it is necessary to engage brain circuits in developing new thoughts and concepts, which depend on already achieved skills, concepts, knowledge, learning, and experiences. Authors express their deep concern that the "made" write-up by ChatGPT will undoubtedly flood the human brain with integrated information but at the cost of helpless and mindless surrendering of thought, reasoning, innovation, executive functions, imagination, novelty seeking, creativity, and writing abilities. Human brain circuits will become lazy by long-term non-utilization. They ultimately will become "sedentary-brain," much like our current lifestyle. Researchers will need to find new ways to counteract the negative effects of a sedentary brain, including the potential rise of a new behavioral addiction: ChatGPT addiction.

#### COMPARISON OF HUMAN AND LARGE LANGUAGE MODELS IN THEORY OF MIND TASKS

This capacity to understand and interpret the mental states of others is referred to as theory of mind.<sup>16</sup> Theory of mind is fundamental to human social interactions, encompassing aspects such as communication, empathy, and decision-making in social contexts.<sup>16</sup> It has been a focal point of research for developmental, social, and clinical psychologists for many years. In a recent study by Strachan et al.,<sup>17</sup> the performance of large language models such as GPT-4, GPT-3.5, and LLaMA2-70B was compared with human participants across various theory of mind tasks. The tasks included understanding false beliefs, interpreting indirect requests, recognizing irony and faux pas, and understanding complex mental states.<sup>17</sup>

- False Belief Task: Both human participants and large language models performed at ceiling levels in the false belief task, demonstrating similar capabilities in understanding where an agent would look for an object based on their memory rather than its current location.
- Irony Task: GPT-4 outperformed humans in recognizing ironic statements, while GPT-3.5 and LLaMA2-70B performed below human levels.
- Faux Pas Task: GPT-4 and GPT-3.5 struggled with detecting faux pas, whereas LLaMA2-70B performed at or above human levels. However, further analysis suggested that the superior performance of LLaMA2-70B might be due to a bias towards attributing ignorance.

- Hinting Task: GPT-4 again showed superior performance compared to humans, while GPT-3.5 was comparable to human performance, and LLaMA2-70B performed worse.
- Strange Stories Task: GPT-4 significantly outperformed humans, whereas GPT-3.5 was on par with humans, and LLaMA2-70B scored lower.

These findings highlight that while large language models like GPT-4 exhibit advanced capabilities in some aspects of the theory of mind, they also have specific areas where their performance diverges from human-like understanding, particularly in tasks requiring nuanced social inferences.<sup>17</sup> The implications of these results are significant for understanding the potential and limitations of large language models in replicating human cognitive processes. The varying performance across different tasks suggests that while large language models can excel in certain cognitive functions, they may still rely on different mechanisms compared to humans.

#### Conclusion

The human brain functions in ways that often go beyond our conscious perception, leaving much of its true potential unacknowledged. Regular activities such as reading and writing play a crucial role in activating numerous brain circuits, thereby enhancing integrated cognitive functions. These activities engage various cognitive domains, including attention, executive function, language, memory, visuospatial abilities, and social cognition.

While artificial intelligence chatbots like ChatGPT offer significant advantages as cognitive enhancers, their impact on long-term cognitive development must be carefully considered. Understanding and anticipating these effects will allow educational and political communities to formulate strategies to mitigate potential adverse outcomes and enhance the positive aspects of artificial intelligence integration into daily life. With the widespread use of ChatGPT, it is essential to recognize the value of maintaining traditional cognitive exercises. Writing, for instance, is not just a language skill but a complex cognitive activity that involves comprehension, knowledge, grammar, visuospatial and perceptual abilities, and imaginative thinking. Engaging in the act of writing helps strengthen these cognitive domains and promotes overall brain health and functionality.

González et al.<sup>18</sup> emphasizes the potential of building a socio-cognitive architecture for collective human-machine intelligence. This architecture

involves planning and coordinating interdependent actions, integrating individual cognitive and metacognitive capacities to achieve common goals. Such integration can enhance collaborative efforts, providing significant societal benefits, including improved decision-making, innovative solutions to complex problems, and enhanced overall productivity.

While ChatGPT offers impressive capabilities and can serve as a valuable tool in various contexts, over-reliance on it for cognitive tasks can lead to the erosion of these essential skills. The ease of generating ready-made content might discourage the practice of writing and critical thinking, leading to potential declines in cognitive abilities. This is particularly concerning for students and individuals in learning environments where the development of these skills is paramount.

We must strive to find a balance between leveraging the advantages of artificial intelligence and preserving our natural cognitive abilities. By continuing to engage in activities that challenge and stimulate our brains, we can ensure that we do not lose the finely tuned cognitive skills that have been developed through centuries of human evolution. Ultimately, while artificial intelligence can aid and enhance our capabilities, it is crucial to remember that the supreme cognitive functions endowed by Nature require continuous practice and engagement. We should embrace artificial intelligence as a tool to complement and support

our cognitive endeavors, not as a replacement for the inherent abilities that define human intelligence.

**Data Statement:** The data generated or analyzed during this study are available from the corresponding author upon reasonable request.

### Author Contributions

All authors contributed significantly to the creation of this manuscript; each fulfilled the criterion established by the ICMJE.

### Acknowledgments

J. Benito-León is supported by the National Institutes of Health, Bethesda, MD, USA (NINDS #R01 NS39422), the European Commission (grant ICT-2011-287739, NeuroTREMOR), and The Recovery, Transformation and Resilience Plan at the Ministry of Science and Innovation (grant TED2021-130174B-C33, NETremor).

**Conflict of Interests:** The authors have no conflicts of interest to declare.

**Study Funding:** This study has not been funded.

**Patent consent:** Not applicable.

**Ethical considerations:** Not applicable as it is a review article.

## References

1. Heyes C. New thinking: the evolution of human cognition. *Philos Trans R Soc Lond B Biol Sci.* 2012 Aug 5;367(1599):2091-6. Doi: 10.1098/rstb.2012.0111. PMID: 22734052; PMCID: PMC3385676.
2. Sachdev PS, Blacker D, Blazer DG, Ganguli M, Jeste DV, Paulsen JS, et al. Classifying neurocognitive disorders: the DSM-5 approach. *Nat Rev Neurol.* 2014 Nov;10(11):634-42. Doi: 10.1038/nrneurol.2014.181. Epub 2014 Sep 30. PMID: 25266297.
3. Beaudoin C, Beauchamp MH. Social cognition. *Handb Clin Neurol.* 2020;173:255-264. Doi: 10.1016/B978-0-444-64150-2.00022-8. PMID: 32958179.
4. Leblanc H, Ramirez S. Linking Social Cognition to Learning and Memory. *J Neurosci.* 2020 Nov 11;40(46):8782-8798. Doi: 10.1523/JNEUROSCI.1280-20.2020. PMID: 33177112; PMCID: PMC7659449.
5. Kärtner J, Schuhmacher N, Giner Torrens M. Culture and early social-cognitive development. *Prog Brain Res.* 2020;254:225-246. Doi: 10.1016/bs.pbr.2020.06.011. Epub 2020 Jul 16. PMID: 32859289.
6. Dowd AJ, Friedlander E, Jonason C, Leer J, Sorensen LZ, Guajardo J, et al. Lifewide Learning for Early Reading Development. *New Dir Child Adolesc Dev.* 2017 Mar;2017(155):31-49. Doi: 10.1002/cad.20193. PMID: 28267289.
7. Freiman M. A 'cognitive turn' in creative writing—Cognition, body and imagination. *New Writing.* 2015;12(2):127-142.
8. Devereux K. Embracing Uncertainty: How literary writing helps us change our minds. *Front Psychol.* 2022 Oct 11;13:1019987. Doi: 10.3389/fpsyg.2022.1019987. PMID: 36304853; PMCID: PMC9595112.
9. O'Sullivan N, Davis P, Billington J, Gonzalez-Diaz V, Corcoran R. "Shall I compare thee": The neural basis of literary awareness, and its benefits to cognition. *Cortex.* 2015 Dec;73:144-57. Doi: 10.1016/j.cortex.2015.08.014. Epub 2015 Aug 28. PMID: 26409018.
10. Gordijn B, Have HT. ChatGPT: evolution or revolution? *Med Health Care Philos.* 2023 Mar;26(1):1-2. Doi: 10.1007/s11019-023-10136-0. PMID: 36656495.
11. The Lancet Digital Health. ChatGPT: friend or foe? *Lancet Digit Health.* 2023 Mar;5(3):e102. Doi: 10.1016/S2589-7500(23)00023-7. Epub 2023 Feb 6. PMID: 36754723.
12. van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. *Nature.* 2023 Feb;614(7947):224-226. Doi: 10.1038/d41586-023-00288-7. PMID: 36737653.
13. León-Domínguez U. Potential cognitive risks of generative transformer-based AI chatbots on higher order executive functions. *Neuropsychology.* 2024 May;38(4):293-308. Doi: 10.1037/neu0000948. Epub 2024 Feb 1. PMID: 38300581.
14. Hagendorff T, Fabi S, Kosinski M. Human-like intuitive behavior and reasoning biases emerged in large language models but disappeared in ChatGPT. *Nat Comput Sci.* 2023 Oct;3(10):833-838. Doi: 10.1038/s43588-023-00527-x. Epub 2023 Oct 5. PMID: 38177754; PMCID: PMC10766525.
15. Breithaupt F, Otenen E, Wright DR, Kruschke JK, Li Y, Tan Y. Humans create more novelty than ChatGPT when asked to retell a story. *Sci Rep.* 2024 Jan 9;14(1):875. Doi: 10.1038/s41598-023-50229-7. PMID: 38195660; PMCID: PMC10776760.
16. Marchetti A, Di Dio C, Cangelosi A, Manzi F, Massaro D. Developing ChatGPT's Theory of Mind. *Front Robot AI.* 2023 May 30;10:1189525. Doi: 10.3389/frobt.2023.1189525. PMID: 37377631; PMCID: PMC10292745.
17. Strachan JWA, Albergo D, Borghini G, Pansardi O, Scaliti E, Gupta S, Saxena K, Rufo A, Panzeri S, Manzi G, Graziano MSA, Becchio C. Testing theory of mind in large language models and humans. *Nat Hum Behav.* 2024 May 20. Doi: 10.1038/s41562-024-01882-z. Epub ahead of print. PMID: 38769463.
18. Gonzalez C, Admoni H, Brown S, Woolley AW. COHUMAN: Building the Socio-Cognitive Architecture of Collective Human-Machine Intelligence. *Top Cogn Sci.* 2023 Jun 18. Doi: 10.1111/tops.12673. Epub ahead of print. PMID: 37331024.