

Comprehensive Analytical Study on Spatial Reasoning and Adaptation of Human Brain Versus Artificial Intelligence - A Review Report

Sahasra V.¹; Yukaciny S.²; Nithiyasri V.³; Santhanalakshmi V.⁴; Logeshwari⁵;
Dr. D. P. Sivasakti Balan⁶; R. J. Thayumanaswamy⁷

^{1,2,3}(Grade 10), ^{4,5}(Mentor), ⁶(Principal), ⁷(CEO)

^{1,2,3,4,5,6,7}KRM Public School, Sembium, Chennai.

Publication Date: 2026/10/17

Abstract: Recent advances in computer vision, spatial computing, robotics, and embodied artificial intelligence (AI) have substantially improved machine capabilities in visual perception, navigation, spatial reasoning, and object manipulation. However, performance on structured benchmarks and controlled environments does not necessarily demonstrate equivalence to human spatial intelligence in complex real-world situations. This review examines whether current AI systems can reproduce the breadth and flexibility of human spatial and embodied intelligence, with particular emphasis on their implications for the potential replacement of humans by AI. The review synthesizes theoretical and empirical literature on human spatial cognition, embodied AI, vision-language-action models, robotics, and relevant AI capability benchmarks. The evidence is organised around four dimensions: (1) biological sensorimotor coupling and proprioception, (2) zero-shot spatial problem-solving and dynamic affordance recognition, (3) physical common-sense reasoning and adaptation, and (4) energy-efficient and context-sensitive spatial processing. Existing research indicates that AI systems have achieved considerable progress in perception, navigation, planning, and manipulation, but continue to encounter challenges involving unfamiliar environments, occlusion, uncertainty, dynamic obstacles, changing object configurations, physical interaction, long-horizon adaptation, and generalisation beyond training distributions. Human spatial intelligence, in contrast, integrates perception, proprioception, embodied experience, contextual knowledge, physical intuition, memory, and flexible real-time decision-making as an interconnected system. The reviewed evidence therefore suggests that current AI can augment and extend human spatial capabilities but cannot yet reproduce the breadth, adaptability, embodied experience, and contextual flexibility of human spatial intelligence. Spatial ability consequently represents an important dimension of human irreplaceability and supports a future paradigm of human-AI complementarity rather than complete human substitution.

Keywords: Artificial Intelligence, Human Intelligence, Spatial Intelligence, Embodied Intelligence, Robotics, Navigation, Manipulation, Spatial Reasoning, Generalisation, Human-AI Collaboration.

How to Cite: Sahasra V.; Yukaciny S.; Nithiyasri V.; Santhanalakshmi V.; Logeshwari; Dr. D. P. Sivasakti Balan; R. J. Thayumanaswamy (2026) Comprehensive Analytical Study on Spatial Reasoning and Adaptation of Human Brain Versus Artificial Intelligence - A Review Report. *International Journal of Innovative Science and Research Technology*, 11(10), 988-997. <https://doi.org/10.38124/ijisrt/26oct455>

I. INTRODUCTION

Artificial Intelligence (AI) has transformed human interaction with technology. Modern AI systems can generate text, recognise images, analyse large datasets, solve mathematical problems, write computer programs, and perform increasingly complex reasoning and planning tasks. These advances raise an important question: if AI can perform many intellectual tasks faster or more accurately than humans, to what extent can humans eventually become replaceable?

This question cannot be assessed solely through digital or language-based tasks. Human intelligence is multidimensional, encompassing reasoning, perception, spatial understanding, physical interaction, adaptation, social understanding, creativity, and decision-making. Among these, spatial and embodied intelligence is particularly important because humans continuously interact with a three-dimensional and dynamic physical world. Everyday activities such as navigating an unfamiliar building or handling a fragile object require the integration of visual perception, depth estimation, memory, proprioception, tactile feedback, physical prediction, and continuous motor adaptation.

Spatial intelligence—the ability to perceive, transform, generate, and recall spatial, visual, and relational information—is fundamental to human interaction with the environment (Korteling et al., 2021). Unlike purely symbolic reasoning, human spatial cognition is closely linked to bodily experience and sensorimotor interaction. This distinction is relevant to Moravec's Paradox, which highlights that tasks requiring high-level reasoning may be relatively easy for computers, while seemingly simple perceptual and sensorimotor tasks remain challenging (Korteling et al., 2021).

Recent advances in computer vision, spatial computing, robotics, and vision-language-action (VLA) models have enabled AI systems to perform increasingly sophisticated spatial tasks, including 3D reconstruction, object recognition, navigation, and manipulation. However, strong performance in structured benchmarks does not necessarily translate into reliable performance in unfamiliar physical environments. Current systems continue to face challenges involving occlusion, uncertainty, dynamic obstacles, changing object configurations, unfamiliar layouts, physical interaction, and adaptation to unexpected conditions (Sun et al., 2024).

The OECD AI Capability Indicators further demonstrate that AI capabilities vary across domains, including vision, manipulation, and robotic intelligence, rather than representing a single general capability. This provides a useful framework for examining whether progress in specific AI tasks corresponds to broader human-equivalent intelligence.

To what extent can current AI systems reproduce the spatial and embodied intelligence that enables humans to function flexibly in unfamiliar physical environments, and what does this reveal about the potential irreplaceability of humans by AI?

The review synthesises research across four dimensions: (1) sensorimotor coupling and proprioception, (2) zero-shot spatial problem-solving and dynamic affordance recognition, (3) physical common-sense reasoning and adaptation, and (4) energy-efficient and context-sensitive spatial processing. By comparing human spatial cognition with contemporary AI and robotic systems, the review evaluates both the progress of AI and the limitations that continue to distinguish human spatial intelligence.

For an experimental study based on this review, the following hypothesis can be proposed: H_1 : Humans will demonstrate greater adaptability and generalisation than AI systems when performing spatial tasks involving unfamiliar environments, occlusion, moving obstacles and changes in object position. Null Hypothesis (H_0) There will be no significant difference between human and AI performance in spatial tasks involving unfamiliar environments, occlusion, moving obstacles and changes in object position. H_2 : AI systems will perform comparatively strongly in structured spatial tasks but will show a greater reduction in performance than humans when task conditions change unexpectedly.

II. REVIEW OF LITERATURE

➤ *Concept of Spatial and Embodied Intelligence*

• *Spatial Intelligence*

Spatial intelligence refers to the ability to perceive, represent, understand, and reason about objects, locations, distances, directions, shapes, and relationships within space. It encompasses several interrelated abilities, including the perception of three-dimensional environments, estimation of distance and depth, understanding of object-object relationships, mental representation of locations, navigation, prediction of movement, interpretation of orientation and perspective, recognition of changes in spatial arrangements, and planning actions within physical constraints.

Spatial cognition is therefore a multidimensional process rather than an isolated ability. It involves acquiring, organising, using, and updating knowledge about spatial environments and supports functions such as navigation, spatial memory, object manipulation, and action planning.

Spatial intelligence is particularly important when examining the limitations of AI in real-world environments. Unlike purely digital tasks, spatially grounded activities require the integration of perception, reasoning, movement, and physical interaction. Humans can continuously update their understanding of an environment when faced with occlusion, unfamiliar layouts, moving obstacles, altered object positions, changing perspectives, and unexpected physical constraints. These abilities connect spatial reasoning with embodied interaction and adaptive action, making spatial intelligence an important dimension for evaluating whether AI can genuinely replace humans in complex real-world situations.

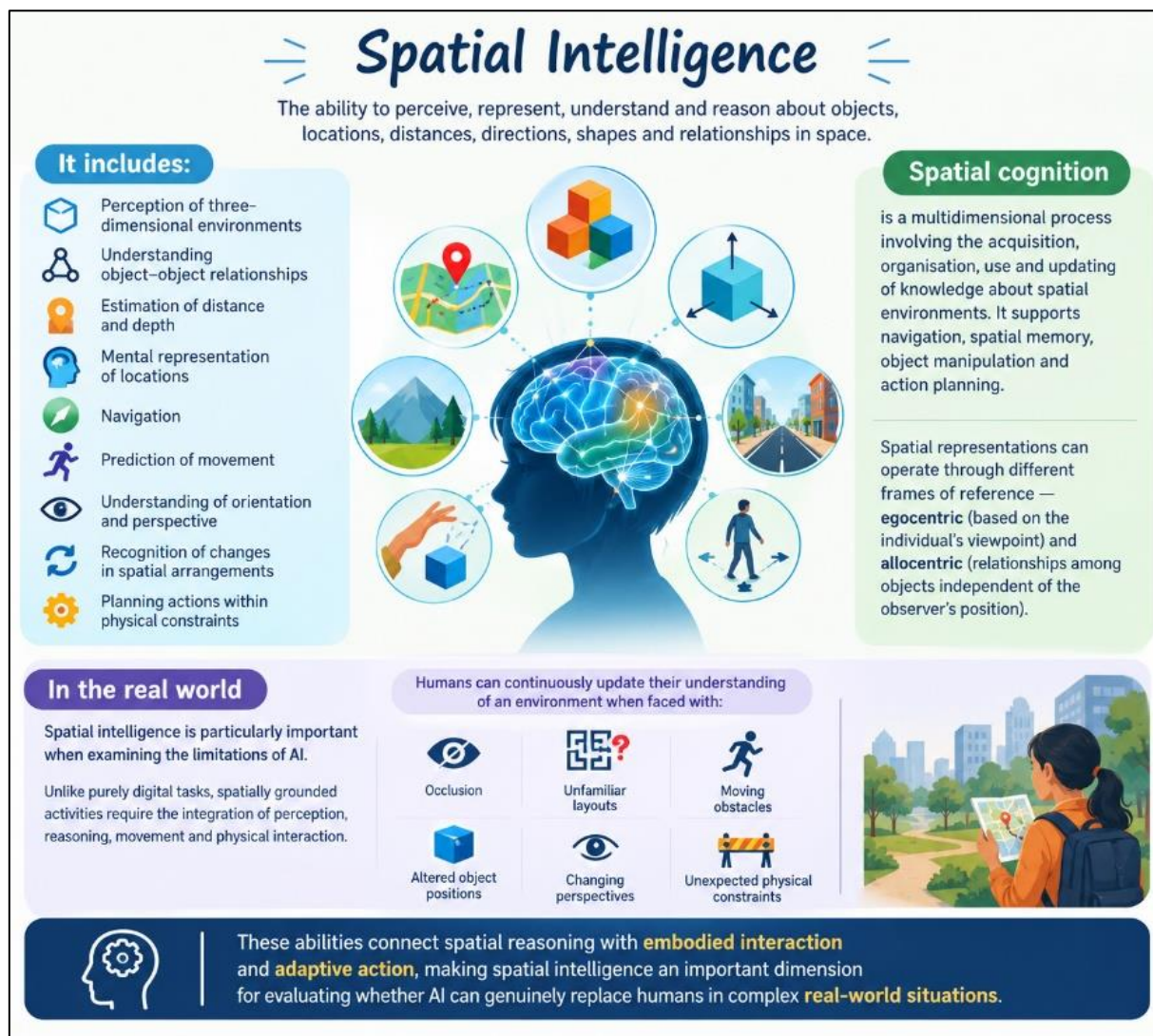


Fig 1 Spatial Intelligence

➤ Embodied Intelligence

Embodied intelligence refers to the ability **to** perceive, understand, and interact with the physical environment through a body, integrating sensory perception, cognition, movement, and action. It involves sensorimotor coordination, object manipulation, force and contact perception, tool use, physical problem-solving, and adaptation to changing conditions. Unlike purely digital intelligence, embodied intelligence requires an agent to continuously use feedback from its body and surroundings to modify its actions.

Embodied intelligence is important when evaluating the potential replaceability of humans by AI because real-world tasks often involve unfamiliar objects, unexpected obstacles, changing physical conditions, failed actions, and real-time interaction. Humans can perceive these changes, adjust their movements, and recover from errors using accumulated physical experience. Therefore, embodied intelligence represents an important dimension for assessing whether AI can move beyond performing predefined digital tasks to reliably and adaptively functioning in complex, unpredictable physical environments.

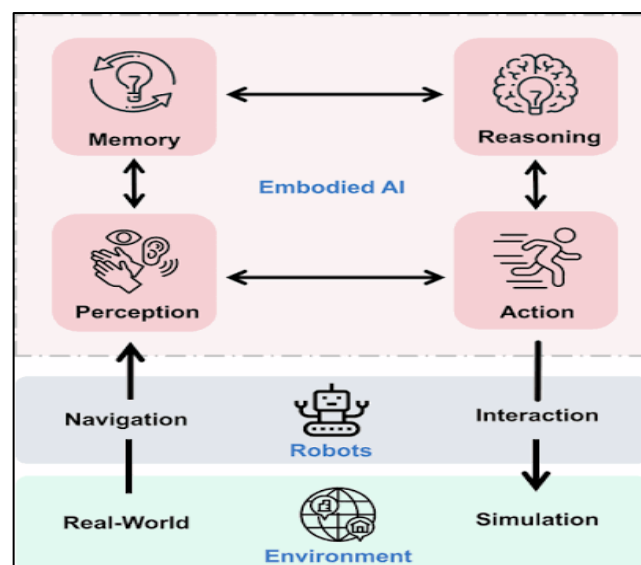


Fig 2 Embodied AI

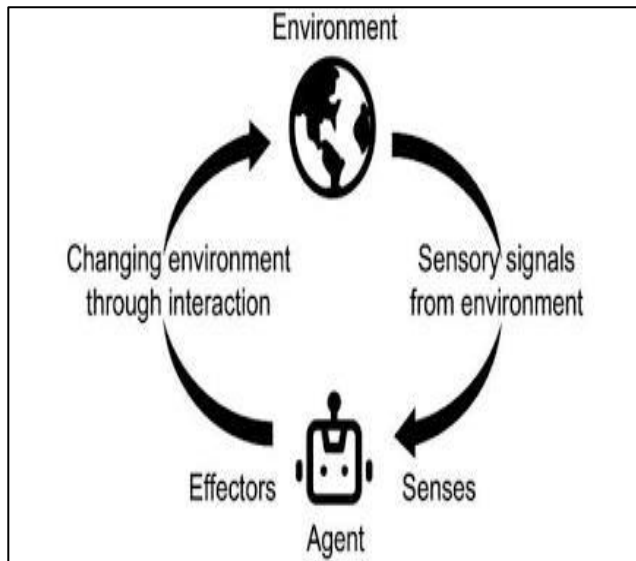


Fig 3 Trasfer of Information

➤ Human Spatial Intelligence

Human spatial intelligence is the ability to perceive, understand, and reason about objects and environments in three-dimensional space. It includes depth perception, navigation, spatial relationships, mental rotation, movement prediction, and adaptation to changing surroundings.

For example, a person walking through an unfamiliar room can judge distances, avoid obstacles, locate objects, and adjust their path when an object is moved. This demonstrates how humans combine visual perception, memory, physical experience, and real-time adaptation—capabilities that remain challenging for AI in unpredictable physical environments.

• Key Human Spatial Capabilities Include:

- ✓ 3D perception: Humans interpret visual information to understand the structure and depth of their surroundings.
- ✓ Spatial relationships: They can determine how objects are positioned relative to one another, such as identifying whether an object is *behind*, *beside*, *above*, or *inside* another.
- ✓ Navigation: Humans can create mental maps and navigate both familiar and unfamiliar environments using landmarks, directions, and spatial memory.
- ✓ Depth and distance estimation: People can judge relative distance, size, and depth and use these judgments to interact safely with their surroundings.
- ✓ Mental rotation and visualization: Humans can mentally rotate objects, visualize different perspectives, and predict how objects will appear or fit when their position changes.
- ✓ Object manipulation: Spatial intelligence supports reaching, grasping, assembling, arranging, and using objects by understanding their shape, orientation, and physical constraints.
- ✓ Adaptation to changing environments: Humans can rapidly adjust their actions when objects move, obstacles appear, layouts change, or expected conditions are disrupted.

- ✓ Reasoning under incomplete information: Humans can infer the location or properties of partially hidden objects using context and previous experience.
- ✓ Embodied interaction: Human spatial reasoning is closely connected with sensory perception, movement, touch, and physical experience, allowing people to understand space while actively interacting with it.



Fig 4 Multimodal Perseption

➤ Multimodal Perception

Multimodal perception is the human ability to integrate information from multiple senses—such as vision, hearing, touch, smell, and proprioception—to understand and respond to the environment. Rather than processing each sensory input independently, humans combine these signals to form a richer and more context-sensitive understanding of situations.

For example, when crossing a road, a person simultaneously sees approaching vehicles, hears their sounds, judges their distance and speed, and senses their own body position, allowing them to make an appropriate decision.

➤ Dynamic Spatial Understanding

- Dynamic spatial understanding refers to the ability to perceive, predict, and respond to changes in the spatial environment over time. Humans continuously track moving objects, changing distances, directions, obstacles, and relationships between objects, while simultaneously adapting their own movements.
- For example, while walking through a crowded corridor, a person can track moving people, anticipate their paths, avoid collisions, and adjust their route when the environment changes.

Spatial intelligence therefore involves continuous updating, rather than simply recognising a scene.

➤ Generalisation

Humans can apply previously learned knowledge to unfamiliar environments and situations, even when they have not encountered them before. This allows them to infer the function of objects, understand spatial relationships, and anticipate possible outcomes without specific prior training.

- *For Example, a Person who has Never Entered a Particular Building can Generally Infer:*

- ✓ A doorway is an entrance;
- ✓ A staircase leads between levels;
- ✓ A chair can obstruct movement;
- ✓ A wet floor may be slippery;
- ✓ A narrow opening may restrict movement.

Thus, human generalisation enables flexible adaptation to novel environments by transferring learned knowledge and principles to situations that have never been directly experienced.

➤ *Physical Common Sense*

Physical common sense is the human ability to understand how objects and the physical world are likely to behave based on everyday experience, even without explicit instructions. It helps humans predict movement, stability, weight, force, collision, and cause-and-effect relationships.

Example: If a person sees a glass placed near the edge of a table, they can anticipate that pushing it further may

cause it to fall and break. Similarly, they know that a heavy box is harder to lift and that an unsupported object may fall.

Thus, physical common sense enables humans to make rapid, context-based predictions about real-world events and adapt their actions accordingly.

Research on physical reasoning has specifically identified the ability to reason about the behaviour of physical objects and choose appropriate actions as a major challenge for AI.

III. METHODOLOGY

➤ *Current Capabilities and Spatial Limitations of Artificial Intelligence*

Current AI systems, particularly computer vision, vision-language models (VLMs), multimodal AI, and embodied robotic systems, have demonstrated substantial progress in spatial intelligence. They can perceive objects, identify spatial relationships, estimate depth, interpret scenes, navigate environments, and perform selected physical manipulation tasks. However, these capabilities remain uneven and are strongly influenced by the complexity and familiarity of the environment.

Table 1 Spatial Dimensions

Spatial Dimension	Current AI Capability	Key Limitation
Spatial perception	Recognises objects, scenes, shapes, positions, and visual features.	Can be affected by clutter, ambiguity, unusual viewpoints, and incomplete information.
3D & depth perception	Estimates depth and constructs 3D representations from visual and sensor data.	Complex depth, perspective, scale, and 3D geometry can remain challenging.
Spatial relationships	Identifies relationships such as left/right, above/below, near/far, and inside/outside.	Multi-object and complex relational reasoning can be inconsistent.
Navigation	Navigates mapped and controlled environments using cameras, GPS, LiDAR, and maps.	Unfamiliar layouts and unexpected environmental changes remain difficult.
Dynamic spatial understanding	Tracks moving objects and changing spatial relationships.	May struggle with relative motion, unpredictable movement, and rapidly changing scenes.
Occlusion handling	Can infer partially hidden objects using visual context.	Performance may decrease when critical spatial information is concealed.
Spatial generalisation	Transfers learned spatial concepts across related tasks and environments.	Generalisation to genuinely novel physical situations remains limited.
Physical common sense	Can learn patterns concerning object position, movement, and interaction.	Lacks human-like intuitive understanding of gravity, weight, friction, stability, and causality.
Object manipulation	Robots can detect, grasp, move, and arrange selected objects.	Irregular objects and unexpected physical conditions can cause failures.
Embodied interaction	Connects perception, reasoning, and action in robotic systems.	Reliable perception-to-action adaptation in open-ended environments remains challenging.
Adaptation & error recovery	Can re-plan actions when explicitly designed for changing conditions.	Humans generally adapt more flexibly to unforeseen changes and failures.

The table demonstrates an important distinction: AI is increasingly capable of performing spatial tasks, but task performance is not equivalent to human-like spatial intelligence. AI can process visual and spatial information at remarkable speed, yet its performance may deteriorate when conditions become unfamiliar, dynamic, physically interactive, or incomplete.

Humans, in contrast, continuously combine perception, spatial memory, physical common sense, prior experience, bodily feedback, and contextual reasoning. This enables them to enter unfamiliar environments, infer hidden information, predict the consequences of physical actions, and modify their behaviour when circumstances change.

Overall, current AI demonstrates substantial competence in spatial perception, 3D scene analysis, spatial

reasoning, navigation, and robotic manipulation. Nevertheless, persistent limitations remain in spatial generalisation, dynamic understanding, occlusion, physical common sense, and adaptive interaction with unpredictable environments. The central limitation is therefore not the inability of AI to process spatial information, but its difficulty in achieving the flexible, embodied, and context-sensitive spatial understanding characteristic of humans. This distinction provides important evidence for examining the continued irreplaceability of human spatial intelligence.

➤ *Occlusion*

Occlusion Handling can be presented as a sub-aspect of spatial intelligence:

Table 2 Human Capabilities and AI Limitations.

Aspect	Human capability	AI limitation
Occlusion	Infers hidden objects using context and experience	May struggle when important spatial information is hidden
Example	Infers a complete chair behind a table	May misidentify its shape, position, or extent
Importance	Supports navigation and interaction in cluttered environments	Can affect reliable spatial reasoning and physical action

➤ *Current Capabilities and Spatial Limitations of Artificial Intelligence*

Place Occlusion Handling after Dynamic Spatial Understanding and before Spatial Generalisation.

Occlusion highlights the difference between visual recognition and genuine spatial understanding. Humans can infer the presence, shape, and position of partially hidden objects using context, prior experience, and physical common sense. Although AI systems can handle many occluded scenes, their reliability may decrease when spatial information is incomplete or highly complex. Thus, occlusion demonstrates that effective spatial intelligence requires not only seeing objects, but also inferring what cannot be directly seen.

➤ *Navigation and Moving Obstacles*

Navigation involves more than identifying the shortest route between two points. In a crowded environment, humans simultaneously consider the destination, obstacles, moving people, available space, speed, collision risk, and changes in the surrounding environment. They continuously predict movement and modify their path according to changing conditions.

Current embodied-AI research indicates that navigation agents still face challenges in robust path planning, collision avoidance, and active exploration in dynamic environments. The 2025 *IndustryNav* benchmark evaluated vision-language models in industrial scenarios containing moving people and objects and reported notable deficiencies in these areas.

• *Inference:*

These findings suggest that spatial intelligence involves more than detecting obstacles or calculating a route. Effective navigation requires continuous perception, prediction, decision-making, and adaptation to changing physical conditions. Thus, recognising an obstacle is not equivalent to

• *Occlusion Handling:*

Occlusion occurs when one object partially or completely blocks another from view. Humans can often infer the hidden portion, approximate the object's position, and understand its spatial relationship with surrounding objects using context, prior experience, and physical common sense. Current AI systems can also infer partially hidden objects, particularly in familiar visual settings, but performance may become less reliable when occlusion is extensive, objects overlap in complex ways, or the environment is unfamiliar. This highlights the importance of contextual and spatial reasoning beyond simple visual recognition.

understanding how to safely interact with a changing environment.

IV. OBJECT RELATIONSHIPS AND MANIPULATION

Object relationships and manipulation are important components of spatial intelligence because they require understanding where objects are, how they relate to one another, and how physical actions can change those relationships.

Humans can naturally determine whether an object is inside, beside, behind, above, or underneath another object and can use this understanding to interact with it. For example, a person reaching for a cup behind a book may first move the book aside, then grasp the cup. This requires understanding the objects' positions, accessibility, physical constraints, and the consequences of an action.

Current AI and robotic systems can identify objects and perform grasping or manipulation in controlled environments, but they may face difficulties with unfamiliar objects, clutter, changing object positions, and unexpected physical interactions.

➤ *Inference:*

Object manipulation demonstrates that spatial intelligence extends beyond recognising where objects are; it requires understanding their relationships, affordances, and physical consequences of interaction. Humans can flexibly modify their actions according to context, whereas AI systems may be less reliable when objects or physical conditions differ from those encountered during training.

➤ *Spatial Memory*

Spatial memory is the ability to encode, store, and retrieve information about locations, directions, distances,

and relationships between objects. It enables humans to remember where objects are located, recognize familiar places, and navigate environments without continuously relying on external instructions. For example, a person can remember the location of a classroom within a school and reach it even when visual signs are absent.

Human spatial memory is closely connected with experience, perception, navigation, and contextual understanding. It also allows people to update their mental representation when objects or routes change.

Spatial memory demonstrates that human spatial intelligence is not limited to immediate perception; it involves maintaining and updating an internal representation of the environment over time. While AI systems can store and retrieve spatial information from maps and learned representations, humans can integrate remembered locations with experience and contextual changes to navigate flexibly in unfamiliar or changing environments.

➤ *Adaptation to Changed Object Positions*

Adaptation to changed object positions is the ability to notice when an object has moved, understand how the new position affects the surrounding space, and adjust one's actions accordingly. It combines spatial perception, spatial memory, and flexible decision-making.

➤ *Example:*

A student usually walks through a classroom using a clear space between two desks. If one desk is unexpectedly moved into that pathway, the student immediately notices the change, understands that the usual route is blocked, and chooses another path without needing explicit instructions.

This demonstrates that humans do not simply remember where objects *were*; they continuously update their mental representation of the environment and adapt their behaviour when spatial relationships change.

Adaptation to changed object positions demonstrates the dynamic nature of human spatial intelligence. Humans can compare current surroundings with stored spatial information, detect changes, reassess spatial relationships, and modify their actions accordingly. This continuous updating and adaptation is particularly important in unpredictable real-world environments and remains a challenge for AI systems.

➤ *Physical Reasoning*

Spatial intelligence becomes particularly significant when combined with physical reasoning—the ability to understand what actions are physically possible within a given environment.

For example, if a system encounters a large box, a narrow doorway, a small table, and a fragile object, it must reason not only about what each object is, but also about their physical properties, spatial relationships, and possible consequences of interaction. It must determine whether the box can pass through the doorway, whether the table provides

sufficient support, and how the fragile object can be moved without damage.

• *Thus, the Question is Not Simply:*

“What is this object?”

✓ *But Rather:*

“What can I do with this object, given its properties and the surrounding environment?”

Physical-reasoning benchmarks have shown that reasoning about physical scenarios and selecting appropriate actions remains challenging for artificial systems.

• *Inference:*

Human spatial intelligence is fundamentally action-oriented: humans continuously combine perception, physical knowledge, and contextual understanding to predict consequences and guide their actions. This ability to reason about what can be done, what may happen, and how to act safely is an important distinction between human embodied intelligence and current AI systems.

➤ *Evidence from Human–Robot Collaboration*

Human–robot collaboration (HRC) provides evidence that effective physical intelligence requires more than perception and task execution. In shared environments, robots must interpret human actions and intentions, predict movement, coordinate tasks, maintain safety, and adapt their behaviour in real time. A 2024 review of human–robot collaboration notes that current systems often rely on simplified models of human intentions and are generally limited to short-term prediction, while accounting for human adaptation remains challenging.

Evidence from systematic reviews also shows that human factors such as trust, cognitive workload, situation awareness, and anxiety can influence collaborative performance, efficiency, response time, and how people use robotic systems. PubMed Central (PMC) More recent HRC research emphasizes adaptive task planning, dynamic role allocation, multimodal communication, and continuous feedback as important requirements for effective collaboration.

• *Example:*

In a shared workspace, a human worker may unexpectedly change direction, reach for an object, or alter the order of a task. A collaborative robot must recognise the change, predict its implications, and adjust its movement or task plan without compromising safety.

Human–robot collaboration demonstrates that physical intelligence is fundamentally interactive and adaptive. Robots can perform increasingly sophisticated tasks, but effective collaboration requires understanding human behaviour, predicting changing intentions, adapting actions, and maintaining safety in shared environments. This supports the view that human intelligence remains particularly important in unpredictable physical contexts, where flexibility,

contextual understanding, and real-time adaptation are essential.

➤ Evidence Does Not Mean AI Is Incapable of Progress

It is important not to overstate the argument. AI and robotics are advancing rapidly, with recent research demonstrating increasingly capable systems for complex manipulation, 3D scene understanding, language interaction, and long-horizon autonomy. One 2026 study reported a 91.67% success rate across 240 real-world manipulation trials for its proposed system. Similarly, recent robotics research

has demonstrated improved generalisation to unseen terrains, payloads, speeds, and external disturbances through physics-informed feedback mechanisms.

• “AI Cannot Develop Spatial Intelligence.”

“AI is progressively developing spatial and embodied capabilities; however, current systems have not yet demonstrated reliable, general, human-equivalent spatial intelligence across the breadth of unfamiliar, dynamic, and physically complex environments.”

➤ Analysis

Table 3 Comparative Analysis: Humans vs AI

Spatial/Embodied Ability	Humans	Current AI/Robotics
Object recognition	Highly flexible	Strong in many domains
Depth perception	Rapid and integrated	Increasingly capable but sensor/model dependent
Spatial relationships	Flexible and contextual	Strong on many benchmarked relationships; generalisation remains challenging
Navigation	Works in familiar and unfamiliar environments	Strong in structured settings; dynamic/unseen environments remain challenging
Occlusion reasoning	Can infer hidden objects	Improving but remains challenging
Moving obstacles	Continuous adaptation	Possible, but reliability varies
Object manipulation	Highly flexible	Strong in controlled tasks; irregular and dynamic situations remain difficult
Tactile interaction	Naturally integrated	Requires specialised sensors and integration
Physical common sense	Developed through embodied experience	Improving but not universally reliable
Spatial memory	Long-term and experience-based	Requires explicit memory/map mechanisms
Unfamiliar environments	Strong generalisation	Major research challenge
Unexpected changes	Rapid behavioural adaptation	Improving, but distribution shift remains a problem
Multi-step physical tasks	Flexible	Increasingly capable but still limited
Human-robot coordination	Natural and context-sensitive	Developing rapidly but remains difficult

The OECD's current assessment reinforces this comparison: its state-of-the-art ratings place vision at Level 3, manipulation at Level 2 and robotic intelligence at Level 2, with Level 5 representing full human-equivalent capability. OECD

➤ Why Spatial Ability Matters to Human Irreplaceability

Spatial ability is a fundamental component of human intelligence because humans must continuously perceive, remember, interpret, and act within physical environments. It integrates spatial perception, memory, navigation, physical reasoning, object relationships, and adaptation to changing conditions. Unlike isolated computational tasks, real-world spatial behaviour requires these abilities to operate simultaneously and respond to incomplete, dynamic, and unpredictable information.

For example, a person walking through an unfamiliar building must remember locations, interpret signs and surroundings, avoid moving obstacles, infer hidden objects, understand physical constraints, and modify their route when conditions change. These processes demonstrate that human spatial intelligence is not simply the ability to recognise where objects are, but the ability to understand the environment and act appropriately within it.

Although AI and robotics are increasingly capable of performing individual spatial tasks, reliable human-equivalent performance across unfamiliar and dynamically changing physical environments has not yet been established. Therefore, spatial ability provides an important perspective on human irreplaceability: human advantage lies not necessarily in performing every spatial task better than AI, but in integrating perception, experience, physical reasoning, adaptability, and action into a flexible understanding of the real world.

➤ Limitations

The literature indicates several important gaps in understanding whether current AI systems can achieve human-equivalent spatial intelligence.

• Gap 1: Benchmark-to-Real-World Gap

Many AI evaluations are conducted in controlled or simulated environments. Real-world environments introduce noise, uncertainty, unexpected changes, and physical constraints that may substantially affect performance.

• Gap 2: Static-to-Dynamic Gap

Recognising and locating objects in a static scene is different from continuously tracking moving objects,

predicting their behaviour, and adapting to changing environmental conditions.

- *Gap 3: Perception-to-Action Gap*

Correctly identifying an object does not necessarily imply the ability to interact with it appropriately. Reliable physical action requires understanding object properties, spatial relationships, and physical consequences.

- *Gap 4: Familiar-to-Unfamiliar Gap*

Strong performance in familiar or previously encountered environments does not necessarily demonstrate the ability to generalise spatial knowledge to genuinely unfamiliar situations.

- *Gap 5: Short-Horizon-to-Long-Horizon Gap*

Completing an individual spatial action does not establish the ability to maintain spatial understanding, plan, monitor progress, and recover from errors across a long sequence of actions.

- *Gap 6: Digital-to-Embodied Gap*

Performance on language, vision, or other digital benchmarks does not fully capture the ability to perceive, reason, and act safely within the physical world.

- *Gap 7: Human–AI Comparison Gap*

More direct comparative studies are needed in which humans and AI perform the same spatial tasks under equivalent conditions, particularly when unfamiliarity, uncertainty, and unexpected environmental changes are introduced.

V. DISCUSSIONS

The evidence reviewed suggests that spatial intelligence is one of the most meaningful dimensions for examining human irreplaceability. AI systems can already recognise objects, interpret scenes, estimate spatial relationships, construct maps, navigate controlled environments, and perform increasingly complex manipulation tasks. These achievements demonstrate substantial progress and should not be underestimated.

However, human spatial intelligence extends beyond recognising or processing spatial information. It represents an integrated capability involving spatial perception, spatial memory, navigation, physical reasoning, object relationships, multimodal perception, generalisation, and adaptation to changing environments. Humans can combine these abilities continuously while interacting with the physical world.

The challenge becomes more significant when environments are unfamiliar, partially observable, dynamic, or physically constrained. Humans can use prior experience and contextual knowledge to infer hidden information, predict the movement of objects and people, modify routes, understand physical consequences, and adapt when previously familiar spatial relationships change.

The OECD AI Capability Indicators further support the importance of distinguishing between different forms of AI capability, particularly between digital capabilities and capabilities involving physical manipulation and robotic interaction. At the same time, recent advances in embodied AI, spatial reasoning, physics-informed learning, and robotic manipulation demonstrate that these limitations are active areas of technological development rather than permanent barriers.

Consequently, the strongest scientific position is not that humans are permanently irreplaceable. Such a conclusion cannot be established from current evidence. Rather, the literature indicates that:

Humans currently demonstrate a broader and more flexible form of spatial and embodied intelligence than current AI systems, particularly when tasks require generalisation, physical interaction, contextual reasoning, and adaptation to unexpected environmental changes.

This distinction is central to understanding human–AI relationships. The evidence does not suggest that AI cannot develop these capabilities; instead, it indicates that current AI performance has not yet demonstrated reliable, general, human-equivalent spatial intelligence across the full complexity of real-world environments.

VI. CONCLUSION

Artificial Intelligence has made substantial progress in spatial perception, navigation, physical reasoning, and robotic manipulation. However, current systems have not yet demonstrated the generality, adaptability, and integration characteristic of human spatial and embodied intelligence. The limitations become particularly evident when AI encounters unfamiliar spatial configurations, occlusion, incomplete information, moving obstacles, changed object positions, physical constraints, irregular objects, dynamic environments, and long-horizon tasks requiring continuous adaptation.

Spatial intelligence therefore provides an important dimension for examining human irreplaceability. The evidence does not suggest that AI is incapable of performing spatial tasks; rather, it indicates that AI can replace humans in specific spatial functions, but current evidence does not establish that AI can replace the full spectrum of human spatial and embodied intelligence.

This conclusion should not be interpreted as a permanent technological boundary. AI and robotics are developing rapidly, and capabilities that are currently limited may improve substantially. Therefore, the most scientifically defensible position is that human irreplaceability is presently associated with the broader integration of perception, experience, physical reasoning, contextual understanding, adaptability, and responsible action.

The future may consequently lie less in AI replacing humans and more in AI augmenting human capabilities. AI

can contribute computational speed, perception, prediction, and planning, while humans continue to contribute flexible adaptation, embodied experience, contextual judgement, and responsibility.

➤ *Ultimately, the Critical Question is Not Simply:*

“Can AI perform a task?”

- *But:*

“Can AI perform that task reliably, safely, and adaptively when the physical world changes in ways it has not previously encountered?”

This distinction provides a more rigorous basis for evaluating the continuing role of humans in an increasingly intelligent technological world.

study investigates robotic generalisation to unseen physical conditions.

- [12.] Gui, S., Gui, K., & Luximon, Y. (2026). *Explicit task reasoning empowering robotic manipulation*. npj Artificial Intelligence. The study examines explicit reasoning for complex robotic manipulation.

REFERENCES

- [1.] OECD. (2025). *Introducing the OECD AI Capability Indicators*. OECD Publishing. The framework evaluates AI across nine human-capability domains, including vision, manipulation and robotic intelligence.
- [2.] OECD. (2025). *OECD AI Capability Indicators Technical Report*. OECD Publishing. The technical report provides the methodology and evidence underlying the capability scales.
- [3.] Fisher, R. B., Cohn, A. G., & Lochhead, C. (2025). Vision scale in *Introducing the OECD AI Capability Indicators*. OECD Publishing.
- [4.] Messina, E. R. (2025). Manipulation scale in *Introducing the OECD AI Capability Indicators*. OECD Publishing.
- [5.] Ho, C., Martin, R., Francis, J., & Oh, J. (2025). Robotic intelligence scale in *Introducing the OECD AI Capability Indicators*. OECD Publishing.
- [6.] Mon-Williams, R. et al. (2025). *Embodied large language models enable robots to complete complex tasks in unpredictable environments*. Nature Machine Intelligence, 7, 592–601.
- [7.] *Phy-Q as a measure for physical reasoning intelligence*. Nature Machine Intelligence. The study evaluates AI systems on physical reasoning and generalisation across physical scenarios.
- [8.] *PARTNR: A Benchmark for Planning and Reasoning in Embodied Multi-Agent Tasks*. ICLR 2025. The benchmark evaluates spatial, temporal and collaborative constraints in human–robot tasks.
- [9.] Raychaudhuri, S., & Chang, A. X. (2025). *Semantic Mapping in Indoor Embodied AI—A Comprehensive Survey and Future Directions*. The study reviews spatial mapping and reasoning challenges in unfamiliar indoor environments.
- [10.] *Brain-inspired spatial intelligence for embodied agents*. Nature Communications (2026). The study examines spatial memory and generalisation as important components of embodied intelligence.
- [11.] Jia, J. et al. (2026). *Physics filtering favors the generalization of robot learning*. npj Robotics. The