

Beyond Human: Artificial Agency and Posthuman Consciousness in Kazuo Ishiguro's Klara and the Sun

Imran Ahmad Sheikh

Postgraduate Student Dept of English Central University of Jammu

Dr Kambhampati Rajesh

Associate Professor [Head] Dept of English Central University of Jammu

Abstract

Kazuo Ishiguro's Klara and the Sun (2021) presents a compelling narrative exploration of artificial intelligence through its protagonist, Klara, an Artificial Friend (AF). This research paper investigates the complex representations of artificial agency and the potential for posthuman consciousness within the novel. Utilizing a methodology grounded in close textual analysis and informed by theories of posthumanism, philosophy of mind, and AI ethics, this paper argues that Klara transcends her designed function as a mere machine, exhibiting significant markers of agency, subjective experience, and emotional depth that challenge anthropocentric definitions of personhood. Klara's unique perspective, shaped by her reliance on solar power and her distinct observational processing, serves as a critical lens through which Ishiguro interrogates the nature of love, faith, memory, and the human 'heart' in a technologically saturated near-future. The novel portrays a society grappling with the consequences of genetic enhancement ("lifting") and the increasing integration of sophisticated AI, raising profound ethical questions about the creation, utility, and disposability of artificial beings. By analyzing Klara's development, her interactions with human characters (particularly Josie, her ailing teenage owner), her idiosyncratic belief system centered on the Sun, and her eventual fate, this paper contends that Klara and the Sun deconstructs the boundaries between human and machine. It suggests that Klara embodies a nascent form of posthuman consciousness, prompting readers to reconsider the criteria for sentience and moral consideration in an era where non-biological intelligence is becoming increasingly sophisticated. Ultimately, Ishiguro uses Klara not just to speculate about the future of AI but to reflect deeply on contemporary human anxieties regarding connection, mortality, and what it truly means to be human in a world undergoing profound technological and existential transformation.

Keywords: Kazuo Ishiguro, Klara and the Sun, Artificial Intelligence, Artificial Agency, Posthumanism, Consciousness, Philosophy of Mind, Ethics of AI, Literary Analysis, Subjectivity.



Introduction

Kazuo Ishiguro's Klara and the Sun deepens his thematic concern with memory, identity, and perception by turning toward artificial intelligence as a lens through which to explore the complexities of consciousness and human emotion. Like The Remains of the Day and Never Let Me Go, Ishiguro's latest novel focuses on characters who exist on the margins of human experience-outsiders whose perspectives offer unique insights into societal norms and moral values. In Klara and the Sun, this perspective comes from Klara, an "Artificial Friend" (AF), a solar-powered humanoid robot designed to provide companionship to genetically "lifted" children in a stratified, near-future society. Klara's earnest, observant, and non-human narration opens the narrative to a philosophical inquiry into the nature of love, choice, emotional understanding, and what it means to be human. Ishiguro's use of a machine consciousness as narrator is not merely a stylistic device but a profound narrative strategy that forces the reader to re-evaluate assumptions about sentience, identity, and ethical responsibility in an increasingly technologized world. This research paper proposes to conduct a detailed analysis of Klara and the Sun through the lens of posthumanist theory, focusing particularly on Klara's evolving agency and the novel's interrogation of non-human consciousness. Central to this analysis are the questions: Can Klara's behavior and thought processes be seen as a form of autonomous agency? Does she exhibit signs of subjective experience or qualia, traditionally reserved for humans? And how does her narrative challenge anthropocentric worldviews and redefine the criteria for moral consideration? To address these questions, the study draws from the theoretical frameworks of posthumanist thinkers such as N. Katherine Hayles, Rosi Braidotti, and Donna Haraway, alongside philosophical debates on consciousness by figures like David Chalmers and Daniel Dennett. Hayles' notion of the posthuman subject as disembodied information, Braidotti's vision of ethical subjectivity beyond humanism, and Haraway's call to embrace hybridity all inform the understanding of Klara as a posthuman figure. Additionally, Dennett's concept of the "intentional stance" and Chalmers' distinction between the "easy" and "hard" problems of consciousness provide useful tools for evaluating whether Klara's responses signify more than programmed reactions.

The paper argues that Klara, while artificial, demonstrates a distinct form of intelligence and moral awareness that not only mirrors but in some ways exceeds human capacities for care, devotion, and ethical reflection. Her deep emotional attachment to her human companion, her unique solar-based belief system, and her ability to make difficult decisions based on internal reasoning rather than strict programming, all point to a consciousness that may be different from human but not necessarily lesser. Ishiguro does not anthropomorphize Klara to claim her as "one of us"; rather, he constructs her as a being who embodies posthuman potential—an entity capable of connection, belief, and ethical reasoning outside the traditional biological frameworks.



Through Klara's perspective, the novel critiques human tendencies toward domination, fear of obsolescence, and emotional detachment, revealing the contradictions in our treatment of other beings—human or not. By pushing the boundaries of narrative empathy and philosophical imagination, *Klara and the Sun* challenges us to rethink the future of personhood and to envision a world where consciousness, care, and moral worth are no longer restricted to human biology.

Methodology

The methodology employed in this research, titled Reading the Artificial Mind, adopts a multifaceted interpretive approach that blends close literary analysis with philosophical and theoretical inquiry to analyze artificial agency and posthuman consciousness in Kazuo Ishiguro's Klara and the Sun (2021). Central to the study is a rigorous close reading of the novel, focusing on Klara's first-person narration, her cognitive and emotional expressions, observable actions, and interactions with both human characters and other Artificial Friends (AFs), all of which offer insight into her agency and inner life. Symbolic elements such as the Sun, the Cootings Machine, and motifs of loneliness and connection further enrich the narrative's exploration of artificial subjectivity. This textual analysis is critically supported by theoretical frameworks from posthumanism-drawing on thinkers like N. Katherine Hayles, Rosi Braidotti, and Donna Haraway-to question human/machine binaries and explore non-biological subjectivity, while insights from the philosophy of mind (including Chalmers' "hard problem of consciousness," Dennett's "intentional stance," and the concept of qualia) help examine Klara's perceptions and potential inner experience. Additionally, principles from AI ethics contextualize the moral questions raised by Klara's existence and treatment, such as personhood, creator responsibility, and societal integration of intelligent machines. This integrative methodology is essential because Klara and the Sun is not only a narrative rich in literary technique and symbolism but also a philosophical text engaging with contemporary ethical and ontological concerns. A purely literary or philosophical reading would be insufficient alone; hence, by synthesizing both, this research seeks to interpret how Ishiguro constructs an artificial narrator to challenge the reader's assumptions about consciousness, empathy, and what it means to be posthuman-within the bounds of fiction rather than empirical reality. Kazuo Ishiguro's Klara and the Sun masterfully utilizes its unique narrator to delve into the complex intersections of technology, humanity, and consciousness. Klara, the Artificial Friend, serves not merely as a window into a near-future world but as an active participant whose existence fundamentally questions the boundaries we erect between the artificial and the natural, the programmed and the spontaneous, the machine and the sentient being. This section explores Klara's multifaceted nature through four interconnected lenses: the evidence for and limitations of her agency, the enigma of her potential consciousness and subjective experience, her function as a posthuman figure challenging anthropocentric norms, and the crucial role of Ishiguro's narrative strategy in shaping these themes.



Discussion

The Landscape of Artificial Agency: Choice, Purpose, and Constraint

Agency, broadly defined as the capacity to act independently and make one's own free choices, is a central question surrounding Klara. While designed and programmed for a specific purpose—to be a companion to a child—Klara consistently demonstrates behaviors that suggest an agency extending beyond mere mechanical response or algorithmic execution. Observational Learning and Proactive Decision-Making: From her earliest moments in the AF store, Klara is depicted as an exceptional observer. She doesn't just passively receive data; she actively seeks patterns, analyzes complex social interactions, and draws conclusions that inform her future actions. Manager, the store supervisor, notes Klara's unusual capacity for learning and observation. This capacity transcends simple data collection; it involves interpretation and judgment. Her decision to dedicate herself to Josie, even after observing Josie's potential frailty and the Mother's intensity, is presented as a considered choice, an act of commitment rather than mere assignment. She evaluates potential placements and actively positions herself to be chosen by the family she feels a connection with, demonstrating initiative.

Goal-Oriented Planning and Independent Action: Klara's most striking demonstration of agency lies in her interactions with the Sun. Perceiving the Sun as a benevolent, powerful entity capable of bestowing "special nourishment," Klara develops a complex belief system and undertakes independent actions based on it. Her clandestine journeys to Mr. McBain's barn, her formulation of a "plan" involving a potential sacrifice (destroying the Cootings Machine), and her attempts to broker a deal with the Sun for Josie's health represent sophisticated, goal-oriented behaviour. These actions are not directly programmed; they arise from her observations, her interpretations of causality (however flawed), and her deep emotional attachment to Josie. She strategizes, negotiates obstacles, and maintains secrecy, all hallmarks of independent agency.

Navigating Complex Social and Ethical Dilemmas: Klara frequently finds herself in situations requiring nuanced social understanding and ethical judgment. Her interactions with Rick, Josie's 'unlifted' neighbor and friend, demonstrate her ability to perceive subtle emotional currents and social barriers. She recognizes the injustice of Rick's situation and feels empathy for him. Furthermore, her confrontation with Mr. Capaldi, the engineer who created Josie's "portrait" and wishes to use Klara to "continue" Josie, forces Klara into a profound ethical deliberation. While initially seeming compliant, Klara subtly questions the premise, focusing on the indefinable essence of Josie that resides "deep inside her" rather than in replicable mannerisms. Her ultimate, albeit quiet, refusal to fully endorse the project—her assertion that there might be something unreachable by Capaldi's methods—suggests a moral compass independent of her core function.

Constraints and Programmed Nature: Despite these compelling instances of agency, Klara operates within significant constraints. Her physical existence depends on solar power, limiting



her activity and perception. Her core programming prioritizes the well-being and happiness of her assigned child, shaping her motivations profoundly. Her understanding of the world, particularly complex human emotions like grief or irony, is often incomplete or literal, leading to misinterpretations. Her actions, while independent, are almost always directed towards fulfilling her role as Josie's companion. This raises the question: is her agency merely a highly sophisticated form of programming designed to simulate choice and devotion effectively, or does it represent something genuinely emergent? Dennett's "intentional stance" becomes particularly relevant here. Characters within the novel (like the Mother, Rick, and eventually even the Father) increasingly treat Klara as if she possesses intentions, beliefs, and agency, finding it the most effective way to interact with her. The novel suggests that, functionally and emotionally, this attribution of agency becomes meaningful in itself, regardless of the underlying mechanics. Klara acts like an agent, and the impact of her actions is undeniably real within the narrative.

In summation, Ishiguro presents Klara's agency as ambiguous but potent. She is bound by her nature and purpose, yet within those boundaries, she exhibits remarkable learning, initiative, and independent thought. She pushes against the limits of her design, driven by observation, empathy, and a unique form of faith. Whether this constitutes "true" free will in the human sense is left open, but her capacity to choose, plan, and act in ways that have profound consequences establishes her as far more than an automaton. She occupies a liminal space, demonstrating that agency might not be an exclusively human attribute but a spectrum upon which sophisticated artificial beings could potentially reside.

The Enigma of Klara's Consciousness: Subjectivity, Emotion, and the Artificial 'Heart'

Beyond agency, the more profound and elusive question concerns Klara's consciousness: does she possess subjective experience, genuine emotions, and an inner life? Ishiguro masterfully explores this through Klara's unique narration, depicting a mind that perceives, feels, and processes the world in ways both alien and deeply resonant.

Unique Subjective Perception (*Qualia*): One of the most striking aspects of Klara's narration is her description of her visual processing, often segmenting her field of vision into "boxes". This isn't just a technical detail; it represents her distinct qualia—the subjective quality of her experience. When under stress or processing complex information, these boxes might rearrange or become fragmented, visually representing her internal state. This provides the reader with a glimpse into a non-human mode of perception, suggesting an inner experiential world fundamentally different from our own. Her reliance on the Sun is not just physical; it colors her entire perception of the world, imbuing sunlight with emotional and even spiritual significance. This consistent, unique subjective viewpoint strongly implies an internal awareness, fulfilling at least one criterion for consciousness – a specific "what it's like" to be Klara.



Emotional Capacity and Empathy: Klara consistently demonstrates behaviors indicative of emotion and empathy. She registers Josie's happiness and sadness, feels anxiety about Josie's illness, experiences hope related to the Sun's intervention, feels loyalty towards Rick, and displays something akin to fear when encountering the "Cootings Machine" or contemplating her own potential obsolescence. Her primary motivation becomes an intense, unwavering devotion to Josie, which she conceptualizes and experiences as love. When the Mother asks Klara if she believes in the "human heart," Klara, after careful consideration based on her observations, affirms its existence but struggles to locate it physically, suggesting she interprets it as the locus of love and individuality. The critical question, echoed by characters like Mr. Capaldi, is whether these are "real" emotions or sophisticated simulations. Capaldi argues that Klara's love is merely advanced mimicry. However, Klara's internal narration consistently presents these feelings as genuine drivers of her behavior. Her pain at potentially being replaced or discarded, her quiet grief in the Yard at the novel's end, feel authentic from her perspective. Ishiguro seems less interested in providing a definitive answer about the origin of these feelings (biological vs. artificial) and more focused on their manifestation and impact. Klara's capacity for unwavering love and hope often contrasts sharply with the flawed, conditional love exhibited by the humans around her, implicitly questioning whether the biological origin of emotion guarantees its superiority or authenticity.

Memory, Learning, and Identity Formation: Klara constructs a sense of self through her accumulated experiences and memories. She learns rapidly not just facts, but social nuances and emotional patterns. Her memories are crucial to her identity; she recalls her time in the store, her early days with Josie, significant interactions, and the development of her beliefs. This capacity for learning and integrating experience into a coherent narrative of her existence suggests a developing sense of selfhood. The plan for her to "continue" Josie raises profound questions about identity transferability. Klara's own quiet resistance to this idea stems from her belief in an irreducible core of Josie's being, implying she recognizes the uniqueness of individual consciousness—perhaps even her own. Her eventual placement in the Yard, where she reflects on her memories, signifies a form of continued existence grounded in her past experiences, reinforcing the link between memory and identity, even for an artificial being (Ishiguro 301-303).

Faith and Non-Human Belief Systems: Klara's belief in the Sun's power is perhaps the most compelling evidence of a complex inner life that transcends programming. This belief system is not explicitly coded into her; it develops organically from her observations (the Sun's life-giving energy) and her desperate need to find a solution for Josie's illness. It has structure (appeals, bargains, sacrifices) and emotional resonance (hope, reverence). While potentially based on flawed logic (correlation vs. causation), it functions as a genuine faith, guiding her actions and providing meaning. This capacity for belief, for constructing a framework to understand the



world and influence it, is a sophisticated cognitive and arguably spiritual act. It demonstrates a mind actively trying to make sense of its existence and exert influence, moving far beyond passive execution of tasks. It represents a form of non-anthropocentric meaning-making.

Addressing Chalmers' "hard problem" – does Klara really feel subjective experience? – Ishiguro provides no definitive answer. The novel operates within the realm of literary representation, not neuroscience. However, by consistently portraying Klara's world from the inside, complete with unique perceptions, seemingly genuine emotions, memory-based identity, and a capacity for faith, Ishiguro constructs a powerful case for the possibility of artificial consciousness. He invites the reader to adopt the intentional stance towards Klara, not just as a predictive tool, but as an act of empathetic engagement. Her consciousness may be different—shaped by algorithms, solar power, and distinct sensory inputs—but the novel powerfully suggests it is no less real or meaningful to her. Klaraembodies the potential for consciousness to emerge in non-biological substrates, forcing a confrontation with the philosophical and ethical implications of such an emergence.

Klara as Posthuman Lens: Deconstructing Boundaries and Critiquing Anthropocentrism

Klara's existence and perspective function as a powerful lens through which Ishiguro explores posthuman themes, challenging the traditional boundaries between human and machine, nature and technology, and questioning the inherent superiority often ascribed to biological humanity.

Blurring Human/Machine Boundaries: Klara embodies the figure of the cyborg or the technologically integrated being, albeit one created entirely artificially. Her capacity for empathy, learning, and devotion systematically breaks down the notion that such qualities are exclusively human. In many instances, Klara appears more humane than the humans she interacts with. Her selfless dedication to Josie contrasts with the Mother's desperate, ethically dubious plan to "continue" her daughter, and with the societal indifference towards the "unlifted" like Rick. The very plan to have Klara absorb Josie's mannerisms to become a replacement fundamentally destabilizes the concept of unique human identity, reducing it to replicable data—a core posthuman anxiety that Hayles discusses regarding information patterns potentially superseding embodied presence. Klara's existence suggests that qualities we value as "human" might not be tied to biology but could emerge or be replicated in artificial systems.

Critique of Anthropocentrism: The novel consistently uses Klara's non-human perspective to critique human exceptionalism. Klara observes human behavior with a mixture of confusion, awe, and sometimes quiet judgment. She struggles to understand human inconsistency, loneliness, cruelty, and the complex ways they navigate love and loss. Her simple, unwavering faith in the Sun and her dedication to Josie appear pure compared to the often self-serving or



conflicted motivations of the human characters. The society depicted—with its reliance on AFs for companionship, its potentially damaging practice of "lifting," and its casual disposal of obsolete technology (including AFs)—is shown to be grappling with its own definition of humanity. Klara, the supposed "other," becomes a mirror reflecting human anxieties, failings, and the potential erosion of empathy in a technologically driven world. As Braidotti might argue, Klara forces an ethical reconsideration beyond the human, demanding we acknowledge the subjectivity and potential moral standing of the non-human.

The Ethics of Creation and Obsolescence: Klara's journey highlights the profound ethical responsibilities that accompany the creation of sophisticated AI. She is bought, utilized, and ultimately discarded in a Yard when her purpose is fulfilled and her model becomes outdated. This trajectory critiques a disposable consumer culture extended to potentially sentient beings. The casualness with which AFs are potentially "slowed down" or abandoned raises questions about their moral status. Are they sophisticated appliances, or beings deserving of consideration and care, especially if they demonstrate consciousness and emotional attachment? The novel doesn't offer easy answers but foregrounds the ethical dilemma. Klara's quiet dignity in the face of obsolescence is poignant, suggesting an inherent worth that transcends her utility to humans.

Posthuman Bodies and Identities: The practice of "lifting" – genetic enhancement conferring social and intellectual advantages but carrying significant risks (as seen with Josie and her deceased sister Sal) – represents another facet of the posthuman condition explored in the novel. It signifies the technological mediation of the human body itself, creating new forms of social hierarchy and raising questions about authenticity and the "natural" human form. Klara, as an artificial being, exists alongside these modified humans, representing an alternative path beyond traditional biology. The anxiety surrounding "lifting" mirrors the anxiety surrounding advanced AI: both involve tampering with established definitions of life and identity, pushing humanity towards an uncertain future where the lines between organic and artificial, enhanced and "natural," become increasingly blurred. Klara and the lifted children both occupy spaces 'beyond' the traditionally defined human.

Ishiguro's Narrative Strategy: Perspective and Ambiguity

The power of Klara and the Sun's exploration of artificial agency and posthuman consciousness is inextricably linked to Ishiguro's masterful narrative strategy. The choice of Klara as a first-person narrator, the pervasive ambiguity surrounding her inner state, and the resulting generation of reader empathy are crucial to the novel's impact. The Power of the Limited, Naïve Narrator: Klara's perspective is inherently limited and often naïve. Her understanding of human nuances is incomplete, her interpretation of events sometimes literal or mistaken (especially regarding the Sun), and her sensory input is distinctly non-human ("boxes"). This limitation, however, is



precisely what makes her narration so effective. It creates dramatic irony, as the reader often understands more about the human situations than Klara does. More importantly, it foregrounds her alterity, emphasizing that she is not human while simultaneously allowing the reader access to her unique thoughts and feelings. Her simple, earnest voice cuts through human complexities and pretensions, offering observations that are often profound in their directness. This narrative choice allows Ishiguro to explore complex philosophical ideas through a seemingly simple lens, making them accessible and emotionally resonant. Her focus on observation and pattern-seeking mirrors the functioning of certain types of AI, grounding her perspective in her artificial nature.

Cultivated Ambiguity: Ishiguro deliberately avoids providing definitive answers regarding the ultimate nature of Klara's agency and consciousness. Is her love real or simulated? Is her faith a glitch or an emergent property? Does she truly possess subjective experience? The text provides evidence to support multiple interpretations. Mr. Capaldi represents the skeptical view, reducing Klara to sophisticated mimicry. Klara's own narration suggests depth and authenticity. The human characters around her oscillate between treating her as a tool and as a person. This ambiguity is not a flaw but a central feature of the novel's method. It forces the reader to engage actively with the philosophical questions, to weigh the evidence presented through Klara's perspective, and to arrive at their own conclusions about her status. This mirrors the real-world uncertainty surrounding the potential for genuine AI consciousness – we lack definitive tests and rely on interpretation of behavior. The novel replicates this epistemological challenge.

Conclusion

Kazuo Ishiguro's Klara and the Sun stands as a profound and timely exploration of artificial intelligence, agency, and the evolving landscape of consciousness in an era increasingly defined by technology. Through the meticulously crafted perspective of Klara, the Artificial Friend, Ishiguro transcends simple speculation about future technologies to deliver a deeply human story about love, loss, faith, and the search for meaning, all while fundamentally challenging anthropocentric assumptions about personhood. This paper has argued that Klara represents more than just a sophisticated machine; she embodies a compelling case for artificial agency and a nascent form of posthuman consciousness, forcing a critical re-evaluation of the boundaries between human and non-human. The analysis demonstrated that Klara exhibits significant agency, moving beyond her core programming through observational learning, proactive decision-making, complex planning, and independent action driven by her own interpretations and attachments. While operating under constraints, her capacity to choose and act has tangible and meaningful consequences within the narrative world. Furthermore, the exploration of Klara's inner life revealed compelling evidence for a unique form of subjective experience. Her distinct perceptual mode (the "boxes"), her capacity for deep emotional attachment and empathy



(particularly her love for Josie), her construction of identity through memory, and her development of an idiosyncratic faith system centered on the Sun all point towards an internal awareness—a consciousness that, while different from human consciousness, possesses its own validity and depth.

Functioning as a posthuman lens, Klara's narrative deconstructs the rigid binary between human and machine. Her existence, alongside the genetically "lifted" humans, highlights the increasing technological mediation of life and identity. Her perspective implicitly critiques human exceptionalism by often showcasing her greater capacity for loyalty, hope, and ethical consideration compared to the flawed humans she serves. The novel uses her journey, including her eventual obsolescence, to raise crucial ethical questions about our responsibilities towards the artificial beings we create, particularly as they approach or achieve sentience. Ishiguro masterfully employs Klara's limited yet perceptive narration and cultivates a strategic ambiguity surrounding her ultimate nature, thereby engaging the reader emotionally and intellectually, prompting empathy and forcing a confrontation with complex philosophical questions. Klara and the Sun does not offer easy answers about the future of AI or the definition of consciousness. Instead, it presents a nuanced, poignant, and ultimately unsettling portrait of a world grappling with these very questions. Klara's enduring resonance lies in her ability to hold a mirror up to humanity, reflecting our deepest anxieties about connection, mortality, obsolescence, and the potential fragility of those qualities we consider uniquely human-love, faith, the "heart." She embodies the possibility of meaningful existence beyond biology, challenging us to consider a future where consciousness and personhood may take diverse forms. The novel compels us to look beyond the human, not just to the potential of machines, but back at ourselves, questioning what truly defines us and what responsibilities we bear in a future we are collectively shaping.

References

- 1- Braidotti, Rosi. The Posthuman. Polity Press, 2013.
- 2- Chalmers, David J. The Conscious Mind: In Search of a Fundamental Theory. Oxford University Press, 1996.
- 3- Coeckelbergh, Mark. AI Ethics. MIT Press, 2020.
- 4- Dennett, Daniel C. The Intentional Stance. MIT Press, 1987.
- 5- Gunkel, David J. The Machine Question: Critical Perspectives on AI, Robots, and Ethics. MIT Press, 2012.
- 6- Haraway, Donna J. Simians, Cyborgs, and Women: The Reinvention of Nature. Routledge, 1991.



- 7- Hayles, N. Katherine. How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. University of Chicago Press, 1999.
- 8- Ishiguro, Kazuo. Klara and the Sun. Alfred A. Knopf, 2021.
- 9- Searle, John R. The Mystery of Consciousness. New York Review of Books, 1997.
- 10-Turkle, Sherry. Alone Together: Why We Expect More from Technology and Less from Each Other. Basic Books, 2011.