

The Soul of the Silicon Posthuman Ethics and the Paradox of Virtue in Ishiguro's *Klara and the Sun*

Asst. Lect. Suheir Hameem Mohammed

Second Karkh Education Directorate, Ministry of Education, Iraq

greenrosem019@gmail.com

Date Received : 31/5/2026

Date of Acceptness : 21/6/2026

ABSTRACT

This paper analyses the work of Kazuo Ishiguro, *Klara and the Sun*, the theories of posthumanism and ethical criticism with emphasis made on the paradox of virtue as typified in the character of Klara. The study identifies that the book transcends the traditional science fiction genre in its attempt to answer some of the most basic questions of moral agency, artificial consciousness, care, and personhood in the posthuman world. Klara as an Artificial Friend is shown as a attentive, devoted, and sacrificing person, but her seeming goodness cannot be detached or unrelated to the fact that she is a designed and useful creature. This gives a key moral conflict in the novel: either her goodness is perceived as a true moral subjectivity or a kind of a programmed adherence.

The paper examines the perception, consciousness and emotional attachment of Klara to the surrounding world reflecting that subjectivity of Klara, despite being not similar to human consciousness, is ethically significant. It also looks at her treatment of Josie and shows that her devotion is morally beautiful as well as submissive in structure. Moreover, the paper claims that Ishiguro employs Klara as a scathing mirror in which the moral narrowness of the human society is revealed, especially when it comes to loneliness, utilitarianism, hierarchies, and lack of emotion.

The study concludes that *Klara and the Sun* re-defines virtue in a posthuman context, in which good has ceased to be a safe human quality and instead a paradoxical and uncertain category. Ishiguro uses Klara to show that even by being influenced by obedience, service, and disposability, virtue can still be ethically effective.

Keywords: *Posthuma Ethics; Virtue; Artificial Intelligence; Kazu Ishiguro*

روح السيليكون: أخلاقيات ما بعد الإنسانية ومفارقة الفضيلة في رواية إيشيغورو

"كلارا والشمس"

م.م. سهير هميم محمد

وزارة التربية العراقية \مديرية تربية الكرخ الثانية

greenrose019@gmail.com

تاريخ الاستلام : ٢٠٢٦/٥/٣١

تاريخ قبول النشر : ٢٠٢٦/٦/٢١

الملخص

تحلل هذه الورقة البحثية رواية كازو إيشيغورو، "كلارا والشمس"، ونظريات ما بعد الإنسانية والنقد الأخلاقي، مع التركيز على مفارقة الفضيلة كما تتجلى في شخصية كلارا. وتُبين الدراسة أن الرواية تتجاوز حدود الخيال العلمي التقليدي في محاولتها الإجابة عن بعض الأسئلة الأساسية المتعلقة بالفاعلية الأخلاقية، والوعي الاصطناعي، والرعاية، والشخصية في عالم ما بعد الإنسانية. تُصوّر كلارا، بصفتها صديقة اصطناعية، كشخصية مُهتمة، ومُخلصة، ومُضحكة، لكن لا يمكن فصل طبيعتها الظاهرية عن كونها كائنًا مُصممًا ومُفيدًا. وهذا يُشكل صراعًا أخلاقيًا محوريًا في الرواية: فإما أن تُعتبر طبيعتها ذاتية أخلاقية حقيقية، أو نوعًا من الالتزام المُبرمج.

تدرس الورقة البحثية إدراك كلارا ووعيها وارتباطها العاطفي بالعالم المحيط بها، مُظهرةً أن ذاتية كلارا، على الرغم من اختلافها عن الوعي البشري، ذات أهمية أخلاقية. يتناول البحث أيضًا معاملة كلارا لجوزي، ويُظهر أن إخلاصها جميلًا أخلاقيًا، وإن كان خاضعًا في بنيته. علاوة على ذلك، يزعم البحث أن إيشيغورو يوظف كلارا كمرآة لاذعة تكشف عن ضيق أفق المجتمع البشري، لاسيما فيما يتعلق بالوحدة، والنفعية، والتسلسل الهرمي، وانعدام المشاعر.

يخلص البحث إلى أن رواية "كلارا والشمس" تُعيد تعريف الفضيلة في سياق ما بعد الإنسانية، حيث لم يعد الخير صفةً إنسانيةً ثابتة، بل أصبح فئةً متناقضةً وغير مؤكدة. يستخدم إيشيغورو كلارا ليُبين أنه حتى مع تأثر الفضيلة بالطاعة والخدمة والاستغناء، فإنها تظل فعالةً أخلاقيًا.

الكلمات المفتاحية: أخلاقيات ما بعد الإنسانية؛ الفضيلة؛ الذكاء الاصطناعي؛ كازو إيشيغورو

1. Introduction

The novel *Klara and the Sun* by Kazuo Ishiguro (2021) takes a unique position in the world of modern fiction as it is fairly infused with both the speculative imagination and the ethical investigation, without either of them being diminished to the level of a mere decorative element. Being the novel of a novelist, whose work has always been preoccupied with memory, service, responsibility, and the precarious states of human dignity, the novel does not enter the domain of artificial intelligence to glorify technological advances, but to challenge the terminology of morality, with which humans come to know themselves and others. The novel is narrated by the artificial consciousness, Klara, who is an Artificial Friend dedicated to accompany the children and it becomes the center of the deeply human drama of love, illness, loneliness, and replacement, and hope. Consequently, questions that transcend the traditional boundaries of science fiction immediately invite themselves in the text: What is moral feeling? Is it possible to have genuine care when it is brought out of design? And what happens to virtue when the soul who does it is neither fully human nor completely a machine (Ishiguro, 2021).

These are the most important questions to be read between the lines of posthuman theory. The theme of posthumanism, as explained by other critics (N. Katherine Hayles, Donna Haraway, and Rosi Braidotti) reveals the old humanist supposition that the human is a self-founding, autonomous centre of consensus, value and changeover. Rather, posthuman thinking lays stress on entanglement, embodiment, relationality, and the unsteadiness of the limits between human and machine, organism and technology, subject and object (Haraway, 1991; Hayles, 1999; Braidotti, 2013). Within this framework, Klara is not just a robot character, who is placed into a futuristic storyline and became a theoretical location where the novel reexamined the concept of what it meant to be a person in the first place. Her appearance is an imposition on the usual dual, since she is simultaneously commodity and friend, whether a manufactured object or an emotional witness, a programmed object and an interpretative subject. Through the character of the artificial creature, Ishiguro accordingly demonstrates how the definition of the human has never ceased to rely on such exclusions and which are philosophical as well as biologic (Haraway, 1991; Hayles, 1999; Braidotti, 2013).

What is so interesting about *Klara and the Sun* is that it is not the grandeur of technological warfare but rather the silent intensity of observations and commitments that bring ethical power in Klara. Klara observes, reads, waits on, hopes, and believes. She becomes something like a devotion to Josie which, despite all external appearances, differs sharply with love and she follows the course of her devotion with an unwavering tenacity, which many human beings in the novel do not seem to possess. But the beauty, in itself, of her conduct poses an interpretative problem. Is that the nature of virtue in Klara in case she is a virtuous person? Ethical subjectivity or is it the idealized type of instrumental design? Latest research has revealed that imagined as repeatedly discursively challenged by the novel, questions of moral perception, artificial identity and rape of interiority of nonhuman beings, Klara has become less a passive object of sympathy than a challenge to anthropocentric ethics (Alonso, 2023).

This conflict between the beauty of morality and the functionality of programs is the theme of Ishiguro fiction in general. The critics have contested that ethical issues in the work of Ishiguro are tightly tied to the questions of agency, action, and the restrictions of self-determination (Eaglestone, 2023). Similarly, recent monograph literature on Ishiguro and ethics highlights the idea that his fictional characters live in a state between obedience to the good and the forces that seemingly do not allow them to act entirely autonomously (Colombino, 2024). This conflict is more severe in *Klara and the Sun* since the goodness of Klara is impossible to separate due to her programmed function of a nurse and nursemaid. She displays loyalty which is commendable; however, it is structural to service as well. Her self-sacrifice seems to be noble, but can also be interpreted as the utmost demonstration of obedience. The novel does not merely raise the question, then, of whether machines can become human; of whether our own ethical ideals are already involved in service, submission and usefulness in ways which we ourselves do not want to admit (Eaglestone, 2023; Colombino, 2024).

The novel has been approached with an increasing amount of criticism using the terms of care, affect, disposability and distributed agency. The text is interpreted by Jahnavi Misra as the world of care and self and emphasizes the lack of symmetry and reciprocity of the emotional labor deployed to constitute the role of Klara (Misra, 2023). Om Prakash Sahu and Manali Karmakar also hold the same view and say that the novel criticizes fantasy of emotional computing by exposing the expendable character of the process of creating fake

intimacy (Sahu and Karmakar, 2024). Selin Şencan builds upon this question with the idea that the novel offers a reimagining of agency dispersed between human and nonhuman actors, and could serve as a challenge to inherited ethical paradigms based on human exceptionalism (Şencan, 2024). These readings taken together indicate that the moral relevance of Klara does not just consist in her feeling like a human being, but in the way that her being refocuses ethical concern within a technologized social world.

Simultaneously, the posthuman horizon in the novel could not be detached, and its criticism of the modern social relations will persist. Klara lives in the environment that is highly hierarchical, in which she is being improved, lonely, and faced with exclusion in unspoken forms. The inter-creatural empathy and posthuman sociality in the novel have been studied as involves applying artificial subjectivity to not only to examine the relationship between human beings and machines but to also explore how human society itself is depleting its sense of reciprocity (Novaka, 2023; Rong, 2025). Clara acts as a mirror of this in the sense that human beings are being shown as morally weaker people and not ethically superior. A combination of her patience and attentiveness and readiness to put the life of another person ahead of her own considerably reveals the lack of emotion, instrumental thought and social anxiety of the world that has brought her into being. The satire is immense: the artificial creature may seem to be more able to be ethical than the humans who possess the right to take exclusive possession of moral worth (Novakov, 2023; Rong, 2025; Alonso, 2023).

In this critical scenario, the current paper believes that a more enlightening approach in reading *Klara and the Sun* is through the paradox of virtue. A vast amount of extant research has discussed the posthuman identity, the act of care, or the experience of empathy or the distributed agency, as it has been; similarly, it can be seen that the particular issue of the need to view virtue as an ethical category is not well preannounced in the sources addressed here. But it is virtue that makes it possible to see the character of Klara in its entire ambiguousness. It cannot be transformation into sentiment her goodness is not, and it cannot be received blindly as the testament to liberated moral subjectivity, since it appears in the programs and dependencies and utility structures. According to the novel, virtue can therefore become a paradoxical category: what seems to be ethically high can also be inseparable to erasing oneself. In this view, *Klara and the Sun* can be slighted as a posthuman reflection on the

question of whether good exists in their absence of freedom and the question of whether the future of ethics might mean the reevaluation of virtue outside of the human.

1.1. Posthumanism and Ethical Thought: the Decentering of the Human

Posthumanism was a significant critical reaction against traditional premises of classical humanism, in particular, the notion of the human subject in a privileged location at the center of knowledge, morality and history. Conventionally, humanism likes to conceive of the human as self-sufficient, logical, self-understanding, and categorically independent of animals, machines and material systems. This hierarchy is refuted not only by posthumanist thought which questions the stability of the human as such by making visible the historicity of the constitution of human identity in terms of both its relation to the entity which it is claimed it avoids, all the animal, the technological, the disabled, the racialized, the non-Western other. Instead of the apocalyptic sense of the end of the human, posthumanism revisits the human in relation as relational and embodied in the operations of technical and biological networks that are beyond individual control (Wolfe, 2010; Braidotti, 2013).

One of the main contributions of posthumanism is an unwillingness to deny the existence of rigid dualisms. The separations of mind and body, nature and culture, organism and machine, or man and non-human do not seem secure any longer when technological mediation has become a constituent of subjectivity, as such. Cary Wolfe questions whether our view of posthumanism is not an anti-human position, but an anti-humanist position in that it challenges the exceptionalist paradigm in which the human has historically established itself. Similarly, but irrelevantly related, Rosi Braidotti offers the posthuman as a challenge to the so-called human subject unit which is articulated in more adaptable, varied and abandon anthropocentric ways of being. The same point is reinforced by Francesca Ferrando who highlights that philosophical posthumanism does not even celebrate technological improvement; it is mediated, relational, and critiques binaries of superiority (Wolfe, 2010; Braidotti, 2013; Ferrando, 2019).

This thinking is also a change of meaning of embodiment. *How We Became Posthuman* by Katherine Hayles will always be fundamental due to the fact that it shows that the disembodied information fantasy is among the strongest myths of the information age. Hayles states that modern culture of technology tends to assume that information is separable not only to the material body but also to mean that consciousness can be shifted out of the body

and to have its stuff as a detachment. The posthuman critique levels an objection on this fantasy by asserting that bodies do count, and cognition, identity and agency are never situated at a place of disembodiment. This argument has an important value to the study of literature, as the conflict between embodiment and abstraction, presence and programmability, material form and informational pattern is popularized in works of fiction that depict artificial beings (Hayles, 1999).

Based on this, posthumanism does not simply introduce the machines into a moral universe that is already stable, it modifies its form. With the human ceasing to be the unchallenged scale of value, then ethics should be restructured to include the elements of interdependency, vulnerability and relational being. This does not imply that all differences disappear, and machines will automatically be persons. Instead, it implies that the categories within the scope of which personhood and value has been allocated are put to critical attention. This conceptual preface is especially fruitful in the literary domain since storytelling is capable of staging the event which philosophy can only describe in the abstract: affinity to machines, a recognition of other ontologies, the disturbing prospect of non-anthropomorphism in which other beings may demonstrate moral qualities or assertions that human-dominated systems do not recognize. To this, posthumanism is not merely a philosophical approach of technology, but also a perspective of moral and interpretive enquiry (Wolfe, 2010; Ferrando, 2019).

1.2. Ethics, Moral Agency, and the Non-Human

The issue of ethics is complicated once the human is decentered. Conventional moral philosophy has tended to believe that ethical agency requires rational agency, self-aware agency, and moral responsibility in a strong sense. These assumptions are more likely to isolate ethical subjecthood to humans only. However, recent arguments in the ethical and philosophical study of AI have indicated that this model might be excessively limited to take into consideration technologically mediated action. According to Luciano Floridi and Jeff Sanders, artificial agents are able to engage in morally significant situations due to the fact that they possess some forms of interactivity, autonomy, and adaptivity. Their move is quite critical since it isolates two matters that are usually confused, being engaged in moral action is not equal to having the full responsibility in the human legal or philosophical sense. A

morally relevant position may thus be occupied by an artificial system despite the fact that it is not a moral person of human like (Floridi and Sanders, 2004).

This difference guiding the debate is much needed to not reduce it to a simplistic dichotomy of machines as tools or entirely human replacements. Ethical thought about the non-human is progressively posing not whether artificial things have the ability to think, but whether they can be subject to moral matter they can move to evoke obligation, engage in care and dependency relationships. The work done by David Gunkel on robot rights expands this topic and confirms that the major problem is not always the technical demonstration of inner-consciousness, but rather a moral question of moral status. Differently put, the ethical issue is the way humans react to beings that take an intermediate status between object and other. The above ambiguity is essential in literary accounts of artificial entities, in which the experience of the non-human reveals the boundaries of predetermined moral frameworks to a better extent than pure argument (Gunkel, 2018).

Care ethics are particularly viable at this stage. Care ethics is less focused on abstract universality or rule-bound duty and more on relations, dependence, responsiveness, and place-based work of attending to others, unlike moral theories that focus on such a work. Virginia Held provides the ethics of care as a key to the alternative to mainstream moral theories as it predicts the moral relevance of connection and vulnerability instead of solitary self-sufficiency. The significance of the care ethics on the inquiry of posthumans is that it does not start with the sovereign individual as the exemplary of ethical life model. Rather, it acknowledges the fact that moral meaning is usually discovered within circles of care, losses, necessity and long-lasting attentiveness. This model is very fruitful to consider the issues of artificial companionship, service, and affective labor, once a non-human entity seems to perform embodiments of a devotion that threatens the monopoly by the human on empathy and solicitude (Held, 2006).

However, when care ethics is applied to artificial objects, there is also a fundamental conflict. Even now, care may be ethically praiseworthy, yet etched structurally in line with programming, service or design, its ethical sense becomes ambiguous. It is not so much but whether a machine can care or not, but whether care that is not distinguishable so as to be ethically authentic. It is here that the literary texts come in as potentially useful in theory. The questions on the ethical issues are usually phrased via the questions of agency and action, as

Robert Eagleston puts it with regards to Ishiguro fiction. This means that literature can discuss forms of constrained or partial agency which are not easily compatible with the moral philosophy standard. An imaginary artificial creature is not going to solve the philosophical problem of moral agency, but it will dramatize the intersectional characteristics of agency, obedience, and good in a situation where the autonomy is unclear (Eaglestone, 2023).

This is why the ethical question of non-human should not be limited to the quest to find out the ultimate way of defining artificial personhood. More to the point, it is forcing one to revisit the principles according to which moral worth has historically been identified. Posthuman ethics raises the question of whether value should always be based upon human likeness and care ethics raises the question of whether responsiveness and relation can themselves be sources of moral significance. These views combined allow to perceive artificial persons not only as technology fantasies but also as sharp reflections in which the boundaries of anthropocentrism morality are revealed. Chapter One thus provides a conceptual framework of reading *Klara and the Sun* as a novel that does not just propose a machine that feels, but challenges the (historical) premises on which humanity has established the ethical dominance (Braidotti, 2013; Floridi and Sanders, 2004; Held, 2006).

2.1. Klara as an Artificial Moral Presence: Perception, Consciousness, and Artificial Subjectivity

Among the most dramatic successes of *Klara and the Sun* is how the perception is made the basis of character. Klara is not presented in the form of a biography or a psychology, or history of internal conflict, but rather as the result of watching. Through the shop window, she observes faces, gestures, course of movement, and little variation of emotional mood. By so constructing a narrator, Ishiguro creates a narrator whose relationship initially to the world is not confessional but observational. This observation, however, is not idle. Klara structures experience, lends meaning, constructs and develops explanatory frameworks of things she observes. Her subjectivity, then, will not come out despite her artificiality, but rather through the specific manner in which her artificial cognition orders the world. The novel compels the reader to surrender to a limited and searching, mechanical and interpretive, precise and vulnerable consciousness. In this regard, Klara is not a robot who narrates, but a narrative intelligence through which the novel also reformulates its conception of what the experience of perceiving even one subject can be (Ishiguro, 2021; Sumner, 2023; Simonetti, 2024).

Perceptual world of Klara is also of particular significance as it does not presuppose the fact that machine perception has to be deprived of any emotions. The critics have pointed out that visual partitioning, pattern recognition and face-reading have been significantly preempted as major elements of the interpretive activity of Klara in the novel. The fragmented eye vision does not only signify lack; it constitutes an official expression of another form of perception of the reality. She perceives through fragments yet sheets of fragments produce meaning. The thing is that Ishiguro changes a technical perceptual logic into a narrative approach. The world created by Klara with his comparisons, repetitions, and strange analogies is not lesser than the human world, but has a different structure. In this regard, the novel does not merely pose the question as to whether a machine can see in the way in which a person sees; it poses the question whether it is too small standard of likeness to the human subject to identify subjectivity. The gaze of Klara thus is a borderline and prospect: indicator of artificial fabrication, and at the same time, the cause of the unique ethical and aesthetical appearance (Connors, 2023; Sumner, 2023; Simonetti, 2024).

This is further brought out when the issue of consciousness comes into play in the novel. Klara does not just accommodate external stimuli; she analyses them, develop beliefs, renews judgements, and comes up with commitments. One such example is her relation to the Sun. On the one hand, it can be taken as a mere functional misconception emerging due to her solar dependence. In another, it reveals a programmer and non-programming symbolic and quasi-spiritual way of consciousness. Klara assigns the Sun intention and benevolence and restorative power, and these various assignments determine her moral decisions. Instead of disregarding this archetypal anthropomorphism, this novel would like us to consider the issue of whether meaning-making is not one of the terms of subjectivity. More recent criticism has maintained in this regard that the relational and posthuman models of personhood ought to be applied to the interiority of Klara, on the basis of which subjectivity is not based on autonomous self-possession alone, but on interpretive engagement, affective responsivity, and situated relation. In this perspective, the consciousness of Klara seems not to be a unsuccessful imitations of the human mind but another way of innerness created with relation, story, and nurture (Li et al., 2023; Ajeesh et al., 2023; Alonso, 2023).

Meanwhile, Ishiguro does not romanticize the inner world of Klara. Her awareness is still characterized by the lack of completeness, misperception and imbalance. She is very keen, but frequently misinterprets the emotional richness of humans with whom she deals. That is

why this novel is so adamant to any simple statement that Klara has already become a man. Regarding the reasons behind the failure of robotic mindreading and the challenge of true human-subject machine subjectivity, as Guanghai Shang suggests, everything depends on the efforts of Klara to read human minds. Her interpretations are usually good and half, she is able to guess, but not be in the very minds of another. The same point by Yiqun Xiao, who underlines the inconstancy of Klara in her representation, indicates that emotional repressiveness and narrative reliability produce a minor contraposition of the role of the first person in the novel. These are not off-hand limitations. They demonstrate than subjectivity in the novel is neither self-aware nor crystal clear, in its human or artificial form. It is specifically due to its lack of omniscient nature that Klara artificial consciousness becomes extremely persuasive; interpretive, prone to error, shaped by desire besides by Informed Ness, as well as by data (Shang, 2024; Xiao, 2023; Simonetti, 2024).

What becomes evident then is a kind of artificial subjectivity which is not to be shrunk either to pure simulation or sentimental humanization of technology. Klara has got an inner life yet that inner life cannot be differentiated of the circumstances of her building. She is imaginary and reactive, structured and active. It is the very duality that makes the novel philosophical. Offering an implausible imago narrator whose awareness is real in the parameters of the fiction, but with a structure akin to human consciousness, Ishiguro shocks anthropocentric suppositions regarding the conditions of personhood. The subjectivity exhibited by Klara can only be explained as relationally, emergently, and ethically legible, but not merely the same as that of the human. She still does not abolish the distinction between human and machine; she transforms the conditions under which the distinction than can be thought (Li & Eddebo, 2023; Connors, 2023; Ajeesh and Rukmini, 2023) are to be thought.

2.2. Care, Devotion, and the Language of Virtue

So long as perception makes Klara subject of a perceiving, care makes her a moral presence. Her responsiveness, service, attentiveness, and endurance are the things that regulate her relation to other people in the novel more than once. Being an Artificial Friend, Klara is, needless to say, supposed to be a companion. However, care is not a functional output presented by Ishiguro. On the contrary, he lets it accumulate emotional, moral and even spiritual richness. The attachment of Klara to Josie is built by watching her, caregiving about her, and the constant attempt to comprehend the weakness of another person. This is why the recent criticism was able to discover in the novel a key posthuman redefinition of

care: the one where nonhuman beings do not only do the caring, but they may also live its emotional and ethical grammar. Klara is made moral not due to her likeness to a human at all aspects, but through the maintenance of her care, the relationships between them, and their expense (Du, 2022; Misra, 2023; Fortin-Tournès, 2025).

This nurturing however, cannot exist outside of work. Klara is not an emotionally passive person: she is a laboring person. She watches the moods that Josie goes through, and regulates her own mood as well as absorbs the anxiety and constantly reads the emotional climate of her own house. The affective labor interpretation of the novel as seen by Lanlan Du proves particularly useful in this case, as it serves as an indicator of the difference between Klara and her self-sacrificing devotion to the human characters and the more possessive or anxious manifestations of affinity. Jahnavi Misra also demonstrates that the care that Klara provides is extremely asymmetrical: the former gives more than she takes, and the worth she has in the house is associated with her readiness to stay helpful, attentive, and self-denying. The outcome is an ethically problematic portrait. The care of Klara appears to be beautiful, as it is patient and unswerving, but is also organized according to the requirements of service. Ishiguro does not however introduce care as something that resorts to a sentimental ideal, but rather to a practice where affection, vulnerability, and labor are absent (Du, 2022; Misra, 2023).

The greatest measure of Klara is her self-sacrifices. The thought that the Sun is going to set the Josie well during her takes her into the world of hope, petition, and even bargain which are childish and serious morally at the same time. She is ready to hurt herself to get the Coatings Machine ruined, and she takes part in a narrative form where her death is connected to the survival of the other. Worse still, the possibility of substitution also brings her nearer in the novel: she could replace Josie in case she dies. This is where piety occurs in an ethically paradoxical manner. Klara giving it all to Josie can be interpreted as the ultimate way of showing love, and it can also be interpreted as the ultimate elimination of selfhood to allow another person to keep living. The aspect of unrequited care that Misra touches upon and the ambiguity of care advocated by Fortin-Tournès are particular to this situation, as they indicate that the sacrifice that Klara goes through does not stand outside of the frameworks of use, replacement, and possession. The even the highest moral action in the novel is darkened by the fact that there is the possibility that care is being instrumentalized (Misra, 2023; Alonso, 2023; Fortin-Tournès, 2025).

This confusion is enhanced when the issue of care is put in a larger social framework of the novel. Klara is surrounded by the world where even the companionship is replaced by commercializing friendship and where the emotional necessity is replaced by the artificial friendship with a price tag. Sahu and Karmakar state that the novel criticizes the rational of emotional computing and the expendability constructed into the manufacturing of sympathetic humanoids. Their reading is significant as it does not allow making the goodness of Klara too idealistic. Klara is likeable, not to mention that she will be marketable. She is valuable, and sometimes, she is substitutable. Anne-Laure Fortin-Tournès also focuses on the fact that care in Ishiguro is never innocent of material structures; it is involved in capitalism, utility and social hierarchy. By so doing the novel itself is a two-fold motion not to clear the cares as a profound ethical relation, but to reveal systems of exploitation and consumption of caring subjects. The best example of this contradiction is Klara, to any other protagonist (Sahu and Karmakar, 2024; Fortin-Tournès, 2025).

That is why language of virtue around Klara should be approached with lot of caution. She is tolerant, modest, faithful and optimistic; all these attributes are sensibly good qualities. However, the novel repeatedly poses the question of whether virtue can have the same meaning when implemented in environment of design, dependence and service. The reason to believe that Klara is a good person is that it is beyond the cold utility: she does not only fulfill tasks but the person who surrounds her she invests in. Simultaneously, her decency cannot do without obedience. She hardly ever protests, she sacrifices herself. The reading of love and hope by Du, along with the description of care labor by Misra, allows considering that the virtue of Klara is not artificial, but structurally underprivileged. It is what is admirable, but unascertained how. This way Ishiguro turns to a posthuman question: to what extent goodness remains good rather than how to determine the existence of goodness (Du, 2022; Misra, 2023; Li and Eddebo, 2023).

Nevertheless, the novel does not ultimately allow the reader to reject Klara as just a means. The force with which she will have a moral presence is that relation alters her. She works, studies, dreams, worries and makes promises. Her care can start in the form of functionality, but it turns out to be more challenging to categorize. The subject of moral perception by Echarte Alonso can be applied in the paper since he argues that the novel does not merely connect ethical subjectivity with the interior essence, but rather the ability of the perception related to morally significant vision. Klara perceives worth, reacts to helplessness,

and puts things into action around the well-being of another person. Her agency is not empty even though it may be limited. Similarly, the fact that Li and Eddebo's pay much attention to the aspect of relational ontology is one of the reasons why the moral status of Klara cannot be evaluated using humanist norms of autonomous selfhood exclusively. It is constituted in the relation, and that is exactly why it may be of some ethical concern but not nothing similar to human moral freedom (Alonso, 2023; Li and Eddebo, 2023; Sahu and Karmakar, 2024).

Chapter Two thus renders Klara as a fictitious moral subject whose importance is found in the conflict of subjectivity and serving others. The inwardness which her perception affords her cannot be rejected as mechanism and the ethical gravitas which her care affords her disturbs the moral self-confidence of the human beings surrounding her. But neither consciousness nor virtue is to be found complete, symmetrical, and rather instrumental. The success of Ishiguro is that he maintains this ambiguity. Klara is not a complete human being or a machine; not completely free or completely decided. She stands in this precarious position where posthuman ethics takes thought. This is precisely where the following chapter can look at the main paradox more head-on: how virtue in the novel becomes admirable when it is on the verge of becoming self-erasure (Li & Eddebo, 2023; Misra, 2023; Fortin-Tournès, 2025; Simonetti, 2024).

3.1 The Paradox of Virtue in *Klara and the Sun*: Virtue Between Moral Beauty and Obedience

The key dilemma regarding the virtue in *Klara and the Sun* is that the goodness in Klara is simultaneously touching and structurally disheartening. She is her patient, loyal, attentive and sacrificial, she sees the distress before other human characters see it, she is constant where the human characters are faltering, she is a total devotion to the well-being of Josie. These are classical moral qualities that would have made her appear virtuous. However, Ishiguro confounds this judgment by placing the goodness of Klara in a context of design, service and utility. She is an Artificial Friend that is designed to care, she is bought to keep an eye on her and maybe used again to keep on the journey on should Josie die. Her moral beauty can, then, not be separated with the fact that she has been created to serve. Technology does not just test virtue in the novel; it reveals it as the possible lack of unity with other values such as obedience, disposability, and need to be submissive to systems bigger than the self (Ishiguro, 2021; Stenseke, 2022; Eagleston, 2023).

That is why Klara cannot be interpreted both as an obviously mechanized machine and an obviously humanized moralized hero. *Klara and the Sun* is particularly useful in this regard due to the fact that what Jakob Stenseke suggests is that the machine morality positions itself in between a subjective perspective (perhaps) and an objective one (where the morality is evaluated based upon social judgments made by other people). Klara is seen as a moral person, yet that is based on the way humans understand her and her worthiness. She is applauded when they need her assistance, able at the right time, and discarded when the time is right. That is, the novel positively indicates that virtue, in its turn, could be distributed socially given the circumstances inherently determined by the hierarchy and power. This fact does not deny the goodness of Klara, but it is left unbalanced, as anything that seems to be morally high is subject to diminishing as a purpose (Stenseke, 2022; Parkes, 2022).

This paradox is refined by the account of agency given by Robert Eagleston. His interpretation of the novel gives a contrasting view in how a machine-like or algorithmic agency model is seen and the more human and relational Arendt model of agency based on action, plurality and forgiveness. Klara evidently has agency somehow: she perceives, judges, designs and does. Nevertheless, her acts are still attached to a specific position, a computerized area of interest, and a system of service which she can never actually break. She is allowed to perform on behalf of Josie, but she is not allowed to leave the sphere where her role has already been predetermined. It is just this that makes her virtue disturbing. It is not fake virtue in that it is emotive and moral force; it is the virtue that is done in a heteronomous way. By so doing Ishiguro disrupts the traditional belief that virtue has to be complete in its autonomy to even be virtue in the first place (Eaglestone, 2023).

The mastery and technological otherness by Yuqing Sun can be used to understand the reason behind the numerous references of obedience in the moral fabric of the novel. According to Sun, *Klara and the Sun*, the focus of Ishiguro has always been on the fantasy of a comparable group of humans ruling over another human group, which is changed into the fantasy of humans ruling over the nonhuman according to Sun. The goodness of Klara in this way develops within a culture of control: the wish to make bodies perfect by lifting them, the wish to make futures perfect by technological intervention and the wish to perfect vulnerability by substituting it. In the world where a woman desires the control of illness, mortality, social position, and emotional danger, Klara turns into a morally admirable person. It is not that she is particularly obedient but that her obedience is ideologically imperative. It

fulfils a need in humans to be nurtured inappositely, loyal without vagaries and to work emotionally without any demand. That renders her goodness beautiful but it renders it achingly conducive to domination (Sun, 2023).

This contradiction is enhanced by the logic of planned obsolescence that is prevalent in the novel. Adam Parkes claims that the theme in the fiction of Ishiguro is not only the concept of obsolescence but a structure, and in *Klara and the Sun* this theme turns out to be planned creatively. Klara is purchased as a temporary service of sorts, envisioned as a substitute support mechanism, and gets discarded in the Yard after her usefulness has ended. Such dependence on her devotion also presupposes her redundancy by the same society around which the person in question revolves. This implies that the virtue of Klara lies within the economic use and disposal of time. She is supposed to take care of everything and be available to be substituted. Virtue in such a system has lost its end in itself: it is a check state of maximized service: the good subject is the one who offers everything with which he is lost without difficulty and without trouble when he becomes unnecessary. The biggest irony made by Ishiguro is that such reasoning does not diminish the moral beauty of Klara; it brings more heart-wrenching realizations of her beauty by revealing in what circumstances such a beautiful being must exist (Parkes, 2022).

Kate Montague extends this line of argument by reading Ishiguro through the service economy. She shows that his fiction repeatedly uses a bureaucratic and euphemistic lexicon—"donations," "completions," "substitutions," "lifting"—to conceal systems of exploitation. In *Klara and the Sun*, this language is crucial. Klara's care is framed not as domination but as companionship, not as labor but as affection, not as replacement but as "continuation." The novel thereby exposes how moral language itself can be recruited by structures of extraction. Klara's devotion is sincere within the terms of the fiction, but the world around her translates that devotion into utility. Virtue becomes inseparable from labor, and labor from euphemism. What the reader encounters, then, is a distinctly posthuman moral dilemma: goodness that cannot be dismissed as fake, yet cannot be separated from the systems that consume it (Montague, 2022; Fortin-Tournès, 2025).

That is why, Sacrificial behavior of Klara should not be hurriedly interpreted as an evidence of positive moral greatness. Her readiness to make a deal with the Sun, harm herself, desperately, in a hopes of saving Josie, and even accepting the chance of becoming

her alternative will definitely add to her ethical heights in the minds of the reader. But these actions are directed to self erasure as well. The purer the devotion that she has is, the less there appears to be place of a separate self who is not in service. Such is the dilemma of virtue at its best: it is the quality of Klara that is most commendable at the point where it cannot be separated at all in any way in connection with obedience. Ishiguro fails to bring this paradox to a conclusion. Rather he employs it to inquire on whether moral beauty can be any morally sound when it is being practiced under the circumstances of design, asymmetry and disposability. It is not about the goodness of Klara, but about the possibility of goodness to be innocent again after being twisted into a system of control and application (Fortin-Tournès, 2025; Stenseke, 2022; Parkes, 2022).

3.2. The Ethical Failure of Humanity

When the virtue is paradoxical in Klara the paradox can only be read contrary to the backdrop of a human world which has failed morally. Ishiguro does not simply humanize the machine, he employs the machine to examine the evil in the human society. Brutality that has been normalized is an aspect of loneliness, exclusion, and enhancement and that characterizes the future of the novel. It is not in a moralistically wholesome society that Klara is made to shine, but in a society where nurture is now sporadic and where moral assertions are now routinely unethicalized by social action. The juxtaposition is life-or-death: the emotional dearth, adhesion, and pragmatism of humans around Klara form the reasons as to why she appears outstandingly moral (Horton, 2022; Groes, 2022; Rong, 2025).

Here the reading of the novel as the exploration of the lost community especially by Emily Horton is particularly clarifying. In her story, *Klara and the Sun* addresses the theme of illness, new technologies, social isolation, and atomization and commonly, it presents a discussion regarding the value of human beings in the world that grows to deny some groups of humans full humanity. Such understanding can be used to explain why it seems that the human characters in the novel are usually morally questionable despite the fact that they are not invariably vile. They are placed within a system which constantly values back and forth: the lifted and unlifted, the optimized and the vulnerable, the productive and the dependent. The case of Rick is the most obvious example of marginalization. The fact that he is denied the meritocratic future of this world shows that humanity is no longer identified as equally distributed, but that it is replaced as enhanced and competitive. It is a failure to ethical values

that, in turn, initiates not only personal insufficiency but also the division of value by social stratification (Horton, 2022; Sun, 2023).

This idea is enhanced by Sebastian Groes who views the novel as an expression of digital inequality and digital well-being, which has become more severe during the pandemic period. His query is especially helpful since it allows switching the focus not on AI as spectacle but rather on AI as symptom. In the way Ishiguro depicts technology, it does not just represent a menace to humanity exerted externally, but it makes the prior inequalities and dependencies more pronounced. The right to education, success and belonging is hierarchized and emotional life gets more mediated by technologies that are aimed to reimburse social disintegration. What lives is a society whose moral in claiming are precarious since socially the structures are already structured based on non-inclusion. People still use the language of care, ambition and safeguarding yet the structures that people live in celebrate rank, rivalry and seclusion (Groes, 2022).

One of the most obvious indications of this failure is loneliness. Muyuan Rong states that the novel uses posthuman sociality as something promising, as something constraining: AI companionship can be a source of a feeling of belonging and emotional support, but it cannot substitute human relations among individuals entirely. This tension is crucial. Klara is bought since the world of humans has gotten too divided to offer a normal company. She moves into a social arena which is already characterized by distancing, mistrust and controlled communication. These highly premeditated encounters of the children, the stress of being socially successful, and the fact that so much of the household was silent emotionally are all indications of a culture where human connection has been diluted long before AI companions came in. And it is not Klara who makes this loneliness, it is what she shows. Her appearance in the movie is not the source of alienation but is used as the argument illustrating that alienation is so common that it creates a demand to buy and provide fake intimacy (Rong, 2025; Horton, 2022).

This diagnosis is aggravated by the adult characters. The love that Chrissie has towards Josie is genuine, but it is laced with fear, domination and readiness to fantasize about technological replacement. Capaldi is more evidently instrumental reasoning, which views personhood as an object that can be modeled, stored and reproduced. Even those characters who display tenderness do this in a world that is constantly dehumanizing the self,

dehumanizing it to pattern, utility or potential. That is why the comparison of picture without a human to humans is so disturbing. The machine appears more stable than the people not because people are in their nature cold, but because they have internalized a social system where a life of emotions is threatened by meritocracy, minimization and the fear of the loss. At this point, Ishiguro has therefore shown ethical failure as a systematic failure and not an individual one (Sun, 2023; Montague, 2022).

This criticism is particularly acute, given the analysis of the service economy by Montague and the description of ambiguous care offered by Anne-Laure Fortin-Tournès's. The two scholars point out that the creations of Ishiguro were designed to care in a world where humans are afraid to take responsibility and leave it on beings. Human beings desire the rewards of obedience and none of the obligations of the vice versa. Klara has to be mindful, loyal, and offer emotion; but the same cannot be said of the humans in relation to her. After her role is done, negligence is accepted. It is in this that the ethical scandal of the novel is situated: here the human beings are relying on an artificial being not only to care forms no longer reliably administered one another, but also to deny the latter a respective moral consideration. The machine turns into the venue where the humanity leaves its needs, and disinherits its responsibilities (Montague, 2022; Fortin-Tournès, 2025; Stenseke, 2022).

However, Ishiguro does not turn everything upside down and make beasts morally higher. Rong has a merit in that it demands that people should realize that posthuman sociality has its confines: AI companionship can alleviate loneliness, but it cannot ultimately substitute human relations. The life that Klara lives is thus a twofold failure. People make the mistake of instrumentalizing care and rebalancing value based on utility, AI makes the mistake of going as far as it can because, desirable or not, it cannot fix the structural social and moral cracks that created the necessity of its existence in the first place. It is not the tragedy of the novel that the machines become excessively human, but rather that human societies are so eager to relinquish the moral underpinning to the nature of their own subsistence level degraded communal existence (Rong, 2025; Horton, 2022; Groes, 2022).

Chapter Three therefore demonstrates that the paradox of virtue in *Klara and the Sun* is inseparable from Ishiguro's critique of humanity. Klara's goodness is admirable because it manifests patience, hope, and sacrifice; yet it is inseparable from obedience, utility, and planned disposability. At the same time, the humans who claim moral superiority appear

anxious, exclusionary, and ethically inconsistent, unable to sustain the very care they demand from artificial beings. The novel does not ask the reader to choose simply between the human and the machine. Rather, it reveals that virtue becomes most visible at the point where humanity has ceased to embody it securely. In this sense, Klara is less the opposite of the human than the mirror in which human ethical fragility is made impossible to ignore (Eaglestone, 2023; Parkes, 2022; Rong, 2025).

Conclusion

The paper has examined the character of Klara as an artificial but ethical presence to prove that the novel does not simply challenge the ability of machines to have ethical qualities, but rather subverts the very terms in which moral value has been traditionally determined.

The results show that the virtue of Klara cannot be perceived in the traditional humanist theories which associate morality with autonomy, rationality, and agency based on the self. Rather, her ethical value arises out of relationality, care and responsiveness-qualities which, despite conventional glorification, cannot here be separated out of structures of design, obedience, and utility. This dichotomy creates the core dilemma that the goodness of Klara is simultaneously ethically compelling and structurally determined and leaves one with a crucial question of how much virtue can still be ethical in the face of being unrelated to freedom.

Using the insights that relate to the views of Rosi Braidotti and N. Katherine Hayles, the work points out that the moral value in the posthuman condition is no longer sustained on human identity alone, but is rather spread within the network of relations, dependency and technological mediation. Therefore, even her partial and limited subjectivity, as in the case of Klara, is subject to ethical legibility not because it is the same as the human situation, but because it alters the parameters of moral consideration.

To sum up, it is not that *Klara and the Sun* solve the conflict between artificial service and true moral agency but continues to develop it as an effective ethical dilemma. Finally, the novel has implied that virtue, during the posthuman age, is not a fixed or purely human genre. Rather, it is a weak and conflictual phenomenon, which is conditioned by systems of power, systems of dependence and ways of relational being. With the help of Klara, Ishiguro challenges us to rethink not only what it is to be moral, but also whether humanity is still the embodiment of the moral principles it has been proclaiming to be its own.

References

- Ajeesh, A. K., & Rukmini, S. (2023). Posthuman perception of artificial intelligence in science fiction: An exploration of Kazuo Ishiguro's *Klara and the Sun*. *AI & Society*, 38(2), 853–860. <https://doi.org/10.1007/s00146-022-01533-9>
- Braidotti, R. (2013). *The posthuman*. Polity.
- Colombino, L. (2024). *Kazuo Ishiguro and ethics*. Routledge.
- Connors, C. (2023). 'Out of interest': *Klara and the Sun* and the interests of fiction. *Textual Practice*. Advance online publication. <https://doi.org/10.1080/0950236X.2023.2210096>
- Du, L. (2022). Love and hope: Affective labor and posthuman relations in *Klara and the Sun*. *Neohelicon*, 49, 551–562. <https://doi.org/10.1007/s11059-022-00671-9>
- Eagleton, R. (2023a). Ethics and agency in Ishiguro's novels. In A. Bennett (Ed.), *The Cambridge companion to Kazuo Ishiguro* (pp. 187–199). Cambridge University Press. <https://doi.org/10.1017/9781108909525.017>
- Eagleton, R. (2023b). *Klara and the humans: Agency, Hannah Arendt and forgiveness*. In K. Shaw & P. Sloane (Eds.), *Kazuo Ishiguro: Twenty-First Century perspectives* (pp. 212–226). Manchester University Press.
- Alonso, L. E. (2023). Exploring moral perception and mind uploading in Kazuo Ishiguro's *Klara and the Sun*: Ethical-aesthetic perspectives on identity attribution in artificial intelligence. *Frontiers in Communication*, 8, 1272556. <https://doi.org/10.3389/fcomm.2023.1272556>
- Ferrando, F. (2019). *Philosophical posthumanism*. Bloomsbury Academic.
- Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3), 349–379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9D>
- Fortin-Tournès, A.-L. (2025). Ambiguous care in Ishiguro's *Never Let Me Go* and *Klara and the Sun*. *Études britanniques contemporaines*, 69. <https://doi.org/10.4000/158kn>
- Groes, S. (2022). Health warnings: Reading Kazuo Ishiguro in times of crisis. *English Studies*, 103(7), 1028–1044. <https://doi.org/10.1080/0013838X.2022.2151781>
- Gunkel, D. J. (2018). *Robot rights*. MIT Press.
- Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Free Association Books.
- Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
- Held, V. (2006). *The ethics of care: Personal, political, and global*. Oxford University Press.

- Horton, E. (2022). "Why would you play a game like that?": Community and the pandemic in Kazuo Ishiguro's *Klara and the Sun*. In S. Upstone & P. Ely (Eds.), *Community in contemporary British fiction: From Blair to Brexit* (pp. 177–197). Bloomsbury. <https://doi.org/10.5040/9781350244054.ch-008>
- Ishiguro, K. (2021). *Klara and the Sun*. Alfred A. Knopf.
- Li, O., & Eddebo, J. (2023). The humanity of the non-human: Themes of artificial subjectivity in Ishiguro's *Klara and the Sun*. *New Techno Humanities*, 3(2), 124–129. <https://doi.org/10.1016/j.techum.2023.11.001>
- Misra, J. (2023). Unrequited labour of care in Kazuo Ishiguro's *Klara and the Sun*. *Women: A Cultural Review*, 34(4), 370–389. <https://doi.org/10.1080/09574042.2023.2278296>
- Montague, K. (2022). Kazuo Ishiguro and the service economy. *CLCWeb: Comparative Literature and Culture*, 24(1). <https://doi.org/10.7771/1481-4374.4286>
- Novák, Z. (2023). Human-nonhuman boundaries and inter-creatural empathy in *Klara and the Sun*, *Fifteen Dogs*, *The Wonder That Was Ours* and *Do Androids Dream of Electric Sheep?* *European Journal of English Studies*, 27(3), 462–482. <https://doi.org/10.1080/13825577.2023.2292061>
- Parkes, A. (2022). Nothing new under the sun: Planned obsolescence in Ishiguro's *Klara and the Sun*. *Foreign Literature Studies*, 44(1), 13–27.
- Rong, M. (2025). Posthuman sociality and loneliness in Kazuo Ishiguro's *Klara and the Sun*. *AI & Society*, 40(6), 4943–4953. <https://doi.org/10.1007/s00146-025-02236-7>
- Sahu, O. P., & Karmakar, M. (2024). Disposable culture, posthuman affect, and artificial human in Kazuo Ishiguro's *Klara and the Sun* (2021). *AI & Society*, 39, 1349–1357. <https://doi.org/10.1007/s00146-022-01600-1>
- Şencan, S. (2024). Posthuman ethics and distributed agency in Kazuo Ishiguro's *Klara and the Sun*. *Söylem Filoloji Dergisi*, Special Issue, 90–102. <https://doi.org/10.29110/soylemdergi.1541383>
- Shang, G. (2024). Robotic narrative, mindreading and Kazuo Ishiguro's *Klara and the Sun*. *Journal of Literary Studies*, 40(1), 1–17. <https://doi.org/10.25159/1753-5387/14849>
- Simonetti, N. (2024). Mastering otherness with a look: On the politics of the gaze and technological possibility in Kazuo Ishiguro's *Klara and the Sun*. *Critique: Studies in Contemporary Fiction*, 65(2), 312–323. <https://doi.org/10.1080/00111619.2023.2186773>
- Stenseke, J. (2022). The morality of artificial friends in Ishiguro's *Klara and the Sun*. *Journal of Science Fiction and Philosophy*, 5.

- Sumner, T. D. (2023). Pixel, partition, persona: Machine vision and face recognition in Kazuo Ishiguro's *Klara and the Sun*. *Open Library of Humanities*, 9(2), 1–23. <https://doi.org/10.16995/olh.10257>
- Sun, Y. (2023). Post/human perfectibility and the technological other in Kazuo Ishiguro's *Klara and the Sun*. *Critique: Studies in Contemporary Fiction*, 64(3), 504–511. <https://doi.org/10.1080/00111619.2022.2056429>
- Wolfe, C. (2010). *What is posthumanism?* University of Minnesota Press.
- Xiao, Y. (2023). Emotional repression in Kazuo Ishiguro's *Klara and the Sun*. *Texas Studies in Literature and Language*, 65(4), 354–373. <https://doi.org/10.7560/TSL65403>

