



Disposable culture, posthuman affect, and artificial human in Kazuo Ishiguro's *Klara and the Sun* (2021)

Om Prakash Sahu¹ · Manali Karmakar²

Received: 12 August 2022 / Accepted: 14 November 2022 / Published online: 27 November 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract

Kazuo Ishiguro's novel *Klara and the Sun* (2021) philosophizes on how in the current technologically saturated culture, the gradual evolution of the empathetic humanoids has, on one hand, problematized our normative notions of cognitive and affective categories, and on the other, has triggered an order of emotional uncanniness due to our reliance on hyperreal real objects for receiving solace and companionship. The novel may be conceived to be a commentary on the emerging discourse in the domain of cognitive and emotional computing that aspires to transform the inner life and social relationships of the human community. The novelty of the paper lies in its ability to showcase how Kazuo Ishiguro's *Klara and the Sun* (2021) creates a rupture in the existing research and literary narrative by critiquing the theoretical underpinnings of emotional computing that optimistically foresees a future where simulated empathetic minds will be able to decode the complexities of the human emotions. It discusses how literature turns into an apt tool to reflect on the limitations of the programmed machines to decode the elusiveness of the human mind that defies the one-to-one correlation between words, multiple connotations, and their underlying emotions. Through the lenses of the fictional narrative, the paper foregrounds how the concept of the social robot designed to offer empathy, care, and companionship turns into a failed project. The paper draws on critical perspectives from disposability theory, posthuman affect, and immaterial bodies to foreground the noncodified feature of affective experientialities that emerge as a result of the interface between humans and nonanimate beings.

Keywords Human–machine interaction (HMI) · Simulated empathy · Affect · Disposability · Posthuman care

1 Introduction

In her 2021 memoir *The Empathy Diaries*, Sherry Turkle philosophizes the computer as an evocative machine that incites self-reflection on a series of fundamental questions pertaining to humans and our essential differences from machines in a technocratic culture where artificial intelligence (AI) is inch-by-inch being programmed to evolve as empathetic minds. The gradual coming into existence of the simulated empathetic machine designed to be a social

companion is the quintessence of hyperreal culture that Turkle argues “offered the illusion of companionship without the demands of friendship and then, as the programs got really good, the illusion of friendship without the demands of intimacy.” (“The Assault on Empathy”, 2018: para 1). Turkle states that “the computer has become an object to think with.” (27). The programmed machine acts almost like a mind. It stands between animate and inanimate, thus evolving as a nonanimate being occupying a space “betwixt and between many worlds” (Turkle 2021: 62). Thus, philosophers and cultural critics argue that in the post-analogue, data-driven world, humanoids are not science fiction characters, they are a part of our social reality, “a cultural other”, turning into an increasingly mature anthropomorphic technological reality (Kim et al. 2009; Kim and Kim 2013: 309; Coeckelberg 2011).

This paper would like to contribute to the ongoing discussion on human–machine interaction (HMI) through the lenses of literary studies. It examines how Kazuo Ishiguro's *Klara and the Sun* (2021) creates a rupture in the existing

✉ Manali Karmakar
manali.karmakar@vit.ac.in; manali.10oct@gmail.com

Om Prakash Sahu
omprakash.sahu@vit.ac.in; omprakashsahu@gmail.com

¹ Vellore Institute of Technology, Chennai Campus, Centre for Innovation and Product Development, School of Electronics Engineering, Chennai, Tamil Nadu, India

² Vellore Institute of Technology, Chennai Campus, School of Social Sciences and Languages, Chennai, Tamil Nadu, India

scientific research narrative by critiquing the theoretical underpinnings of emotional computing that optimistically foresees a future where simulated empathetic humanoids will be able to develop an empathetic relationship with their clients by decoding the complexities of human emotions. This paper through the lenses of literary narrative aims to reflect on the feasibility of reducing human affective orders into algorithmic models that could be configured in a robot for manufacturing social companions. The paper interprets *Klara and the Sun* as a cultural artifact that turns into an apt tool to philosophize on the emotional ambiguities triggered by the increasing social presence of the simulated empathetic minds. The paper argues that Kazuo Ishiguro's fictional description of an empathetic programmed machine serves as a refractive medium enabling humans to introspect on how "transformative technology can change our inner life and relationship" (Turkle 2021: 2). The argument is also supported by Davis and Thacker (1998) who state that artificial intelligence has emerged as an interactive mirror that enables us to "both recognize ourselves in and measure ourselves against it" (cited in Kim and Kim 2013: 310). Kim and Kim (2013) propose six major views to analyze our relationship with a robot as a cultural other. The views are—robot as the frightening other, robot as the subhuman other, robot as the sentient other, robot as the divine other, and robot as the co-evolutionary path to immortality. The sixth view analyses the concept of digital resurrection.

The paper argues that in Ishiguro's *Klara and the Sun*, the proposed views are not demonstrated as starkly opposed to each other, rather the fictional narrative through the HMI demonstrates how the six views are complexly enmeshed in each other. This paper through the analysis of Ishiguro's fictional narrative would like to showcase how the simulated mind emerges as a "non-self" that triggers an "uncanny visceral affect" in humans (Kim and Kim 2013: 310; Ravetto-Biagioli 2019: 59). The notion of the uncanny has its root in Sigmund Freud (2003: 9). Freud explains,

The uncanny is not a direct reaction to mechanical devices but a neurological (internalized) form of automaticity that is realized as an aesthetic or affective experience – *deja vu*, the appearance of a *doppelgänger*, delusions of grandeur, paranoid behaviors, and so on. The uncanny is driven by the compulsion to repeat, and that compulsion is automatic. Rather than a simple reaction to technology, the uncanny is an intentional, embodied (even if symptomatic) response to 'something which ought to have remained hidden but has come to light. (Cited in Ravetto-Biagioli 2016: 1)

It should be noted that Freud's explanation of the uncanny has gone through a phase of metamorphosis in

the post-analogue era. Ravetto-Biagioli (2019) has used the term "digital uncanny" to "explore how digital technologies...are transforming the meaning of the uncanny that Freud tied to the return of the repressed memories" (4). The notion of digital uncanny adds a new dimension to the concept of uncanniness by questioning whether the response we receive from quasi-human objects is subjective or data-driven. The notion of digital uncanny foregrounds the irresolvable uncertainties that generate in humans' an order of intensified affect.

The first section of the paper analyses the notion of posthuman care as dramatized in *Klara and the Sun*. The second half of the paper aims to interpret Ishiguro's novel as a critique of the theoretical underpinnings of emotional computing that optimistically foresees a posthuman future where simulated empathetic minds will be able to offer care and solace to humans. To support its arguments, the paper would specifically focus on the narration of the artificial intelligence named Klara to foreground how the complexities and ambiguities of human emotions stand incomprehensible to a simulated nonhuman, thereby leading to a failure to develop empathetic bonding between humans and machines. It juxtaposes Klara's social skills with the elusiveness of the human mind that defies the one-to-one correlation between words, meanings, and their underlying emotions. Through the lenses of the fictional narrative, the paper foregrounds how the concept of the social robot designed to offer empathy, care, and companionship turns into a failed project. The paper would like to draw on critical perspectives from disposability theory, posthuman affect, and immaterial bodies to foreground the noncodified feature of affective experientialities that emerge as a result of the interface between humans and vital matters.

2 Posthuman care and disposable culture

The novel *Klara and Sun* begins at a store that sells affective robots termed artificial friends (AF) as companions to lonely children who have minimal opportunity to interact with their parents, and to make friends due to 24 × 7 work culture, fragmentation of family unit, and digitalization of education. It may be argued that Ishiguro's novel is an implied reflection on the post-COVID-19 pandemic culture where digitalization and work-from-home (WFH) practices have turned into a new normal. The novel primarily revolves around the life of the fourteen-year-old Josie, a feeble child diagnosed with chronic illness, and hence

in need of constant care and attention, Chrissie, Josie’s mother, Rick, Josie’s intimate friend, and the AF Klara, the narrator. Klara is programmed with outstanding observational qualities that enable her to offer minute details of a society immersed in a technocratic culture where fundamental existential crises such as loneliness, and death are promised to be circumvented by social robots.¹ *Klara and the Sun* has primarily been interpreted as a literary narrative that enables us to reflect on how the notion of being a human is problematized in the technologically driven consumer culture (Lombardo 2021; Askew 2021). Lombardo’s (2021) review paper speculates on the age-old phobia about humans being dethroned by AI. Drawing on the narratives of *Klara and the Sun*, Lombardo argues, “the prospects of truly human-like AI even if it is for now only a pipe dream, rattles that foundation. If our intelligence is all that defines us, who are we when AI matches it?” (110). In a similar vein, Askew (2021) through the fictional narrative investigates what it means to be a human with a specific focus on the bleak prospects of a technologically driven society. Askew primarily foregrounds how humans are gradually being replaced by intelligent machines and genetically upgraded humans popularly known as “lifted”, thus augmenting the social division (Ishiguro 2021: 213). In contrast to the antagonistic explanation of AI, the recent article by Ajeesh and Rukmini (2022) discusses how the science fiction *Klara and the Sun* plays a significant role to reduce apprehensive attitudes, humans have toward social robots. The article argues that the novel celebrates the posthumanist and transhumanist notions of intelligence, care, and empathy that are not confined solely to the human community.

This paper argues that Ishiguro’s novel serves as a prism to sensitize readers to the spectrum of emotional complexities pertaining to human relationships in the rapidly changing posthuman world through the lenses of Klara. Klara, the store-bought, humble, and loyal humanoid robotic caretaker designed to establish empathetic bonding with her client raises ethical questions on the one hand, about the disposable feature of the caring machines, and on the other, enables us to reflect on the underlying theoretical flaws pertaining to emotional computing. The paper argues that the socio-cultural status of the AFs as dramatized in the novel drives readers’ attention to the integral connection between the notion of disposability and affective robots that are conceived as “ontologically hollowed objects” (Kennedy 2008: 144). Greg Kennedy argues that “disposable commodities appear as already disposed of. Their revelation presupposes

their disappearances in the commodified order of technology” (143–144). Ishiguro’s fictional narrative speculates on the ethical issues generated by the concretization of robotics culture where care is translated into “disposable affective things” designed to fulfill the emotional requirements of the affluent section of society (DeFalco 2020: 36). In the novel, the disposable status of the affective robots is aptly captured in the numerous episodes. From the beginning, the fear of the older B2 models being replaced by the upgraded models is captured in the narrative of Klara. Klara narrates “AFs weren’t embarrassed but were afraid. They were afraid because we were new models, and they feared that before long their children would decide it was time to have them thrown away, to be replaced by AFs like us” (15). A similar fate is also experienced by Klara after her client’s requirements are completed. Towards the end of the novel, Klara turns into a redundant object occupying a place firstly in the storeroom and later in the junkyard, “all alone awaiting the end of her lifespan” (Lombardo 2021: 114).

Drawing on Kennedy, it may be argued that AFs like Klara are “carefree commodities” (122). Kennedy argues that in the current culture of consumerism, our constant reliance on technology has led to the production of carefree disposable commodities whose consumption does not require taking care of them. The attitude of carefreeness embedded in the consumer culture facilitates the discarding of commodities inconsiderately after they reach their expiry date. Empathetic humanoids like Klara are designed for a stipulated time frame after the expiry of which they are converted into trash. The carefree attitude towards Klara is aptly demonstrated in the “interaction meeting” episode where Josie offered the AF as a playful object to her friends to assert the rightfulness of her decision to select a B2 model (64). The episode was baffling for Klara as well who is reduced to a mere command-following machine designed to entertain Josie and her visitors. Danny, Josie’s friend, states that it is not morally wrong to swing and throw an AF playfully because “they’re designed to deal with it” (74). The carefree commodity status of Klara is also captured in the episode where Rick’s mother Miss Helen conceived the AF equivalent to a vacuum cleaner, thus proclaiming that the physical presence of Klara does not require paying attention. Miss Helen expresses her views thus, “One never knows how to greet a guest like you. After all, are you a guest at all? Or do I treat you like a vacuum cleaner? I suppose I did as much just now” (145).

The equivalence of Klara with a carefree commodity designed for a specified purpose, nonfulfillment of which may lead to the disposability of the product is further captured in the threatening tone of the Housekeeper Melania thus: “And AF. Your big plan. If it makes Miss Josie worse I come dismantle you. Shove you in the garbage”. (178). The episode also enables us to reflect on how the ontological

¹ The social robots are also termed as carebots that are used at hospitals and homes to assist in, support, and provide care to vulnerable individuals such as the elderly, disabled, and young children.

binary between the human and the artifact is being recurrently manifested in the linguistic construction of the artificial other (Coeckelberg 2011). In the post-analogue society, we are witnessing a transition from talking about a robot to talking to and with a robot. Our perception of the social relationship with the humanoid is influenced by the culture in which we are embedded. Ishiguro by demonstrating the reconstruction of the binary man–machine binary showcases how in the Western world:

Most of us hold to individualistic and non-relational ontologies. Robots are assigned to the object categories, humans to the subject category. No crossovers or hybrids are allowed in this modern, dualist ontology; purity is preserved. Languages already ‘contain’ such ontologies in their structures. Our linguistic grammar is also a moral, social and ontological ‘grammar’. (65)

A robot is an engineering construct but its existence as an artifact or a quasi-other is defined by the socio-cultural constructs that are used to address them. Coeckelberg argues that the language we use allows us to relegate different connotations to the socio-cultural position of a robot. The usage of the second person “you” while addressing a robot creates a hybridized order of social ontology that evolves as a result of the HMI. In this context, the robot occupies the position of a quasi-other, whereas, the usage of “it” reveals the robot as an artifact. However, in *Klara and the Sun*, the position of the social robot Klara problematizes the philosophy of linguistic expression as proposed by Coeckelberg. Through the analysis of the fictional narrative, we are able to deduce that in spite of the fact that Klara is normatively addressed as “you”, her significance as an artifact is not reduced. Klara’s disposable position may further be related to Kennedy’s explanation of the notion of “ontological violence” (144). Kennedy argues that “ontological violence is the progenitor of trash” and the act of not acknowledging the physicality of a being in itself is an act of ontological violence (144). Kennedy states:

Violence negates the physicality of the targeted being. This is most apparent in cases of human being. To do violence to a person means to deny that the person feels pain, suffers, cares for his existence, is mortal. The seeming paradox here is only seeming, for to fully identify the mortality of a human is to necessarily identify with it. In such a state of identity, violence against another has suicidal implications for the aggressor, who can perpetrate it only on the basis of a misunderstanding of himself and the victim. One can harm another when one does not discover oneself in the other. Violence, in other words, depends on not truly discovering the being of the other, not perceiving a person as a person. (118)

Drawing on Kennedy’s explanation it may be argued that the reified and disposable status of Klara presents her as a nonperson. Josie does not acknowledge the affective engagement Klara aspires to establish. The interaction meeting episode aptly captures how Josie reduces her relationship with Klara to a master–slave relationship by ordering the AF to abide by the commands. Klara narrates Josie’s insensitive and authoritative tone thus:

Go on Klara’, Josie said. ‘Go say hello to those boys.

I didn’t move at once, partly because I’d been surprised by Josie’s voice. It was like the one she sometimes used when talking to Melania Housekeeper, but not like any voice she’d used before to me. (75)

The unempathetic attitude as demonstrated by the characters in Ishiguro’s novel reiterates DeFalco’s argument that the etymology of robot is informed by the notion of slave labour and hence “it is no surprise that robotic care is poised to step into the minoritized breach of contemporary work” (35). The notion of drawing a similarity between a robot and a slave is also supported by Bryson (2010). Bryson states, “robots should not be described as persons, nor given legal nor moral responsibility for their actions. Robots are fully owned by us” (1). The reification of Klara further aligns with Clough’s (2007) explanation of how the ideology of “superexploitability” informs our usage of affective robots (cited in DeFalco 2021: 39). Clough argues that inequality is inherent in the concept of affective economies. “Some bodies or bodily capacities are derogated, making their affectivity superexploitable or exhaustible unto death, while other bodies or body capacities collect the value produced through this derogation and exploitation” (2007: 25–26).

In *Klara and the Sun*, the interaction meeting episode has multiple connotations for the readers. On the one hand, it demonstrates the reinforcement of the man–machine hierarchy, and on the other, it manifests how “Josie’s ability to change” her approach towards Klara in selected contexts stands incomprehensible to the AF, thus sensitizing her to the complexities of the human emotions and her limitation to decode those (83). Klara is a data-driven machine, hence there is a one-to-one correlation between the words she communicates and their underlying meaning. Klara is designed to be Josie’s best friend and hence, the situation baffles the AF when Josie draws an equivalence between Klara and the housekeeper Melania. However, in the context of Josie, her mother, and the other characters there is always a discrepancy between the words articulated, multiple connotations, and their underlying emotions (Koshelev 2020; Abdullah 2016). Thus, Klara experiences pragmatic let-downs leading to her failure to develop an empathetic bonding with the humans.

3 Emotional computing and posthuman affect

Zhenhua Zhou (2021) foregrounds the major flaw in the theoretical underpinning of computing-based AI that has hindered the evolvement of machines with real human-like intelligence. Zhou argues that the epistemology of AI is regulated by the “doctrine of cognitive computing” (12). The doctrine believes in the possibility of converting human intelligence into a binary algorithmic convention, and cognizes “human intelligence as the calculation of symbols of mental representation and AI as essentially a computing-based simulation of human thinking” (113). Zhou argues that computing-based AI has failed to realize the fact that “real human intelligence runs through the whole process of the emotions. The core and motivation of rational thinking are derived from emotions. Intelligence without emotion neither exists nor is meaningful” (113). Hence, Zhou emphasizes the development of emotional computing to circumvent the limitations of the existing AI machines.

However, a quick review of the research in the domain of AI and machine learning from 2012 till date foregrounds a contrasting research narrative. Researchers of robotics have been emphasizing the significance of emotional computing for developing affective robots with real human-like intelligence (Wieseke et al. 2012; Destephe et al. 2013; Kirandziska et al. 2014; Kugurakora et al. 2015; Gomez-Martinez et al. 2018; Spatola and Wudarczyk 2021; Corina Pelau et al. 2021; Kerruish 2021). Disciplines such as psychology, biology, anthropology, neuroscience, social science, robotics, and computer science are referred to design algorithmic codes for simulating human-like emotional intelligence. Malinowska (2021) too asserts, that since empathy is conceived to be one of the significant emotional components “allowing people to form and maintain satisfying social relationships with other subjects, it is no surprise that it is one of the most studied phenomena in the area of HRI (Human–Robot Interaction)”. For instance, Samani and Saadatian (2012) claim to have designed an affective robot named “lovotics with the ability to love and be loved by humans” (1). Lovotics employ a probabilistic mathematical model for the formulation of love and empathy. An artificial endocrine system is implemented in the robot by imitating human endocrine functionality, thus the robot is claimed to have the capability of experiencing complex human-like biological and emotional states, as governed by the artificial hormone within the system.

In contrast to the affirmative claims, a Senior Research Associate at the Responsible Technology Institute, in the Department of Computer Science at Oxford University, Dr. Pericle Salvini (2015) sceptically argues:

In Interpersonal relations, social interaction means the capacity to be friendly and pleasant to people. Applied to a robot, the adjective “social” implies several kinds of abilities, among which, cognitive and physical skills as well as social and emotional capacities, such as behaving according to social norms or expressing and understanding emotions and feelings. However, interacting with human beings is a very complex and dynamic activity which can be hardly foreseen and translated into lines of codes. (p. 70)

Scepticism embodied in the argument of Salvini finds resonance in the research findings of Parviainen and Coeckelbergh (2021) who draw on the humanoid robot Sophia designed by the company named Hanson Robotics to express concern about the limitations of simulated embodied intelligence to cognize the complexities of the human emotions. They argue,

Even if robots such as Sophia are equipped with AI (in particular machine learning) that gives it facial recognition and natural language processing capabilities, their movement and robotic “embodied intelligence”, are based on programming to simulate simple everyday tasks. This requires the creation of a four-dimensional (time–space) representation to calculate its movement trajectories so that the robot can perform movements. Due to the necessity of representation, robot movement can never be intentional like human or animal kinaesthesia and motor intentionality

Malinowska adds to the epistemic limitation of the emotional computational theory by foregrounding the flaw in the methodology for cognizing the complexities of emotional components such as empathy. Malinowska states that normatively HRI researchers rely on the psychological definition of emotion that believes in studying affective phenomena through scales, charts, and tables. Maj (2007) and Thompson (2014) criticize this reductive approach to studying the complexities of the affective phenomenon. The application of the psychological definition in the context of HRI is fallible because firstly it offers a very reduced interpretation of the affective complexities. Secondly, emotional aspects such as “empathizing are often considered to be a process correlated with social cognition and social emotions. In the case of robots, it is problematic because social relations are traditionally understood as an interaction of two conscious, intentional, and relational entities” (363). Thus, Malinowska asks, “The most basic doubt is whether the use of social terms and explanations in relation to these inanimate machines is methodologically justified” (363). Hence, it is argued

that in the context of HRI, interdisciplinary research is the urgent requirement for understanding the range of denotative and connotative explanations underlying the affective phenomenon.

This paper concurs with arguments made by Salvini, Parviainen, and Coeckelbergh, and Malinowska. The paper further substantiates the above arguments by referring to theories pertaining to the philosophy of affect to foreground the noncodified nature of human emotions as captured in *Klara and the Sun*. Ishiguro in *Klara and the Sun* draws readers' attention to the limitations of AI to comprehend human emotions from the beginning of the narrative. Being a minute observant, Klara aspires to learn about the spectrum of human emotions by standing at the front window, where she gets the "benefit of an unobstructed view of the street and of the flow of urban life. She is exposed to humans displaying the widest array of emotions, including those bizarre ones like melancholy that seem to be a mixture of opposites" (Lomberdo 2021: 111). In one such incident when Klara was observing urban life, she acknowledges her limitations to comprehend the ambiguities of human emotions thus:

Still, there were other things we saw from the window—other kinds of emotions I didn't at first understand—of which I did eventually find some versions in myself, even if they were perhaps like the shadows made across the floor by the ceiling lamps after the grid went down. There was for instance, what happened with the Coffee Cup Lady. (19)

Ambiguities underlying human emotional expressions and language are aptly expressed in the Coffee Cup Lady episode as well, where Klara with the other AFs stands bewildered witnessing the union of long-lost lovers. The old man and the woman's union manifest a complex array of emotions that defeat the definite labelling of the expressions. The conversation between Klara and the manager of the store dramatizes how the AFs' curiosity to understand the complexities of human emotions and the store manager's effort to explain the underlying significance of the incident stand elusive to the data-driven affective minds:

'Those people seem so pleased to see each other,' Manager said.
And I realized she'd been watching them as closely as I had.
'Yes, they seem so happy,' I said. 'But it's strange because they also seem upset.'...
'Do you mean, Manager, that they lost each other?'...
Manager's voice wasn't the usual one...
'Sometimes,' she said, 'at special moments like that, people feel a pain alongside their happiness... Then Manager was gone, and Rosa said, "How strange. What could she have meant?" (21)

In this context, we may refer to the study of Mower et al. (2009) that foregrounds the underlying differences between human emotions and emotional computation. Mower et al. argue that human emotional experiences are dynamic and are informed by the sociocultural context in which the interlocutors are embedded. Emotional expression is a dynamic interactional process "involving dependencies based on time, speaker, context, mood, personality, and culture" (n.pag). Emotional computing is conventionally based on proto-typical emotional expressions that fail to comprehend the complex range of mixed emotional manifestations "that vary temporally with speech and are expressed and perceived over multiple modalities, may be inherently ambiguous, or may have emotional connotations resulting from other emotional utterances within a dialogue" (n.pag). Mower et al. study is also supported by Quiroga-Clara (2003) who reiterates the limitations of the data-driven machines to decode ambiguities that are an integral component of non-proto-typical emotional expressions.

In alliance with the epistemological problems of defining how emotions can be represented in a robot and under what conditions robots are able to participate effectively in emotional and empathetic dynamics with human beings, we ought to reflect on how the human-techno-world interface termed as "hyperreality" brings significant alteration in the emotional dynamics of humans that further problematizes the codification of the affective phenomenon. The statement corroborates the research findings of Kate Letheran et al. (2020) who underlie the change in human emotions as a result of the increase in interaction with social robots that are programmed to enact empathy with a degree of agency and autonomy (Chapter 3, *Emotions and Service in the Digital Age*). Salvini (2015) argues that the HMI interaction produces an illusion of reciprocity since the robot "will always be responding according to its program and therefore the interaction will always be unidirectional". Salvini explains the illusion of reciprocity as the simulacrum of presence that compels humans to enter a phase of willing suspension of disbelief thus triggering an order of emotional uncanniness. In *Klara and the Sun*, the notion of willing suspension of disbelief is aptly dramatized in the "Morgan's Falls" episode where Chrissie asks Klara to mimic Josie to fill the gap of her daughter who could not accompany her due to ill health (91). Morgan's Falls episode foregrounds the long-term emotional requirement of Chrissie that was not revealed in the initial phase of the novel. Josie is chronically ill and hence, Chrissie purchased Klara to get solace in the hyperreal object that is programmed to seamlessly copy the "impulses and desires" of her daughter (210). In fact, Chrissie aspires Klara to be a digital resurrection of Josie which is revealed in the Mr. Capaldi building episode thus:

Feeling sorry's not what I'm asking of you. I am asking you to do what's within your power... But I promise you I'll never love him the way I'll love you. You'll be Josie and I will always love you over everything else. So do it for me. I'm asking you to do it for me. Continue Josie for me. Come on. Say something. (213)

Morgan's Falls episode turns into a trial phase. The success of the data-driven machine to be the "simulacrum of presence" of Josie fills Chrissie with an uncanny feeling of fear, sadness, and elation that is captured in the observation of Klara:

The Mother leaned closer over the tabletop and her eyes narrowed till her face filled eight boxes, leaving only the peripheral boxes for the waterfall, and for a moment it felt to me her expression varied between one box and the next. In one, for instance, her eyes were laughing cruelly, but in the next they were filled with tears... The Mother leaned even further across the table, and I could see joy, fear, sadness, and laughter in the boxes. Because everything else had gone silent, I could hear repeating under breath: 'That's good, that's good, that's good.' (104).

Chrissie's uncanny expressions underline the vulnerable aspect of humans who in this hyperreal world consciously engage in self-delusional activities with the hope to circumvent biological and socio-cultural limitations. Chrissie's ambiguous reciprocation to the humanoid robot corroborates with the concept of the "uncanny valley" first explained by the Japanese roboticist Masahiro Mori (Mori 1970, 2005; cited in Misselhorn 2008: 345). Morio explains:

The more positive and empathetic emotional responses from human beings it will elicit. However, when a certain degree of likeness is reached, this function is interrupted brusquely, and responses, all of a sudden, become very repulsive. The function only begins to rise again when the object in question becomes almost indistinguishable from real humans.

The uncanny valley is the metaphor used to explain the zone where the binaries between known and unknown, abhorrence and sublimity are problematized by the presence of the humanoid. Misselhorn argues that the feeling of uncanniness is triggered by the disruption in the process of establishing an empathetic relationship with a simulated mind. Empathy is normatively considered to be an affective phenomenon that emerges in the context of humans who are an integral part of social ontology. Social participation occurs with an understanding of shared et the interlocutor "is alive and is endowed with a soul in the Aristotelian sense of the term" (357). However, at the stage where the realization occurs that we are engaging with a hyperreal entity,

a rupture occurs in our feelings due to the quick oscillation in our understanding that the entity which seems "alive and soulful a moment ago now appears cold and dead. The alternation between these two states amounts to the feeling of eeriness" (357).

In *Klara and the Sun*, Klara's presence as a social robot in Josie's family also upholds the notion of willing suspension of disbelief that underlies humans' interaction with simulated minds because we are aware of the fact that "behind the appearance of genuine presence and reciprocity, there is a programme, lines of codes written by another human beings, according to the mathematical model of human behavior which will never be able to predict the complexity and unpredictability of human emotion" (Salvani, 73). In *Klara and the Sun*, Morgan's Fall episode captures how Chrissie's self-delusional act triggers an intensified order of enigmatic experientiality that aligns with the theoretical concept of affect. Affect is explicated as a complex and highly disputable phenomenon that has been approached through multiple perspectives by scholars belonging to the field of psychology, neuroscience, philosophy, and humanities (Ott 2017). In the field of psychology and neuroscience, affect is conceived as an elemental state, whereas philosophy and humanities cognize affect as an intensive force. As an intensive force, the phenomenon is theorized as noncodified, intractable, non-representational, and asignifying order of experientialities by affect theorists such as Deleuze (1987), Massumi (1995), and Bennet (2004).

Affect may be defined as an offshoot of posthuman culture that blurs the dichotomy between terror and pleasure, relief and grief. Vermeulan (2014) states that the assemblages of enigmatic affective experiences have brought "the demise of the strictly codified, subjective feelings", thereby leading to the generation of second-order feelings that defeat any form of codification, definition, and reterritorialization onto the subject. Affective experiences are conceived as second-order feelings because "they emerge in the wake of first-order emotions". In the hyperreal world, engineers and researchers of robotics claim that everything can be translated into data and codes. It is claimed that there are no essential features in humans "that's unique and won't transfer" (201). Ishiguro responds to the claim by dramatizing a contrasting narrative where Klara turns out to be a failed project.

Ishiguro's novel corroborates as well as stands in contrast to popular cinematic representations such as *Ex Machina* (dir. Alex Garland, 2014) that dramatizes the endless possibilities introduced by the evolution of the simulated conscious mind. In the movie, Nathan, a genius entrepreneur and programmer engineers a female humanoid robot Ava. He later conducts a Turing Test with the help of Caleb, an intelligent employee of his company to examine the degree of consciousness as manifested by Ava. However, as the movie gains momentum the binary

between the examiner and the examinee is blurred. Ava, the female humanoid robot has the ability to read humans' expressions better in comparison to humans' reading machines. Caleb fails to offer an appropriate reading of Ava's expression, whereas Ava has access to an archive of millions of human expressions (including Caleb) that enable her to measure and correlate with the real-time facial expressions to establish the genuineness of the expressions as manifested by Caleb. Ava has an advantage over Caleb, and hence, at the end of the test, Caleb encounters an existential crisis. "Caleb's interaction with Ava lead him to question not its humanity but his own." Ex Machina asserts the possibility of "reducing one's subjectivity to data patterns and using them to design objects that would elicit one's genuinely subjective (but effectively preset) responses" (5).

In *Ex Machina*, Ava turns out to be a successful project. Ishiguro negates such possibilities. In the Mr. Capaldi building episode Mr. Paul, Josie's father abhors the notion of creating a digital copy of his chronically ill daughter. With disgust and abhorrence, Mr. Paul asks Klara:

Do you believe in the human heart? I don't mean simply the organ, obviously. I'm speaking in the poetic sense. The human heart. Do you think there is such a thing? Something that makes each of us special and individual?... And that could be difficult, no? Something beyond even your wonderful capabilities. Because an impersonation wouldn't do, however skillful. You'd have to learn her heart and learn it fully, or you'll never become Josie in any sense that matters. (218–219)

Klara responded optimistically by stating, "Josie's heart may well resemble a strange house with rooms inside rooms. But if this were the best way to save Josie... I believe there's a good chance I'd be able to succeed" (219). However, as each day goes by, Klara realizes that there is "something" that makes humans unique (306). Towards the end of the novel in the "Yard" episode during her conversation with the manager of the store, Klara states that the uniqueness of being a human does not lie in the anatomy, rather it is a distributed phenomenon that evolves throughout the life course with the shared emotional engagement with the kith and kin that triggers an order of feelings that make each individual irreplaceable (302). Klara finally concludes the limitations of AI thus:

Mr. Capaldi believed there was nothing special inside Josie that couldn't be continued. He told the Mother he'd searched and searched and found nothing like that. But I believe now he was searching in the wrong place. There was something very special, but it wasn't inside Josie. It was inside those who loved

her. That's why I think now Mr. Capaldi was wrong and I wouldn't have succeeded. (306)

It may be argued that Ishiguro's *Klara and Sun* upholds a posthuman perspective to foreground the essential differences between affective minds and emotional computing. The uniqueness of what it means to be a human is not confined to the biological frame subjected to study under a microscope to offer a taxonomical explanation of the neurobiological underpinning of the human body that could be copied to manufacture simulated nonanimated artifacts. Humanness as an embodied phenomenon evolves as a result of the enmeshment of boundaries between bodies, animals, technologies, spaces, and materialities of different kinds, thus enabling the evolution of "a complex and contradictory mixture of anxieties, desires, and exhilaration" that problematizes the notion of emotions associated with the traditional understanding of human subjectivity (Vermulan 2014: 122).

The domain of robotics has referred to a range of disciplines such as philosophy, psychology, biology, neuroscience, and computer science to cognize the complexities of human emotional factors. This paper would like to conclude its argument by stating that in addition to the stated disciplines, the domain of cognitive and emotional computing may refer to literary narratives that may be comprehended as "case studies or thought experiments" to reflect on the underlying theoretical flaws in the disciplines of Engineering, Science, and Social Sciences that propose to reduce the complexities of emotions into codes for designing data-driven empathetic beings (Robinson, 2009). Ishiguro's *Klara and the Sun* emerges as a critique of this epistemic framework of cognitive and emotional computing that aspires to convert the complexities of affective phenomenon into codes for manufacturing artificial friends.

Funding We have not received funds for this research.

Declarations

Competing interests There is no conflict of interest among the authors.

References

- Abdullah K (2016) Linguistic Arbitrariness According to Saussure. Beltz Verlag
- Ajeesh A, Rukmini S (2022) Posthuman perception of artificial intelligence in science fiction: an exploration of Kazuo Ishiguro's *Klara and the Sun*. *AI & Society*.
- Askew, Rie Kido. (2021). [Review of the book *Klara and the Sun* by Kazuo Ishiguro], 181-185, <https://www.ritsumei.ac.jp/acd/re/ssrc/result/memoirs/kiyou43/43-05.pdf>
- Bennett J (2004) The force of things: steps toward an ecology of matter. *Political Theory* 32(3):347–372

- Bryson J (2010) Robots should be slaves. *Close Engagements with Artificial Companions*. <https://doi.org/10.1075/nlp.8.11bry>
- Clough P (2007) Introduction. In: Clough PT, Halley J (eds) *The affective turn: theorizing the social*. Duke University Press, Durham, pp 1–33
- Coeckelbergh M (2011) You, robot: on the linguistic construction of artificial others. *AI & Soc* 26(1):61–69. <https://doi.org/10.1007/s00146-010-0289-z>
- Davis E, Thacker E (1998) *TechGnosis: Myth, Magic, and Mysticism in the Age of Information* (Reprint ed.). North Atlantic Books
- DeFalco A (2020) Towards a theory of posthuman care: real humans and caring robots. *Body Soc* 26(3):31–60. <https://doi.org/10.1177/1357034x20917450>
- Deleuze G, Guattari F (1987) *A thousand plateaus Capitalism and schizophrenia* (B. Massumi, Trans.). University of Minnesota Press, Minneapolis
- Destephe M, Maruyama T, Zecca M, Hashimoto K, Takanishi A. (2013) Improving the human-robot interaction through emotive movements A special case: Walking. *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. <https://doi.org/10.1109/hri.2013.6483528>
- Freud S, McLintock D, Haughton H (2003). *The Uncanny* (Penguin Classics) (Illustrated ed.). Penguin Classics
- Garland, Alex (Director). (2014). *Ex Machina* (Film). DNA Films.
- Gómez-Martínez DG, Rosales JH, Muñoz-Jiménez V, Ramos F, Ramos M (2018) A bio-inspired self-responding emotional behavior system for virtual creatures. *Biol Inspired Cogn Arch* 26:26–40. <https://doi.org/10.1016/j.bica.2018.07.012>
- Ishiguro K (2022) *Klara and the Sun* (Main ed.). Faber And Faber.
- Kennedy G (2008) *An Ontology of Trash: The Disposable and Its Problematic Nature* (S U N Y Series in Environmental Philosophy and Ethics). State University of New York Press.
- Kerruish E (2021) Assembling human empathy towards care robots: the human labor of robot sociality. *Emot Space Soc* 41:100840. <https://doi.org/10.1016/j.emospa.2021.100840>
- Kim MS, Kim EJ (2013) Humanoid robots as “The Cultural Other”: are we able to love our creations? *AI & Soc* 28(3):309–318. <https://doi.org/10.1007/s00146-012-0397-z>
- Kim MS, Sur J, Gong L (2009) Humans and humanoid social robots in communication contexts. *AI & Soc* 24(4):317–325. <https://doi.org/10.1007/s00146-009-0224-3>
- Kirandziska V, Ackovska N (2014) Comparison of emotion evaluation perception using human voice signals of robots and humans. *Int J Bus Technol* 2(2):13–17. <https://doi.org/10.33107/ijbte.2014.2.2.03>
- Koshelev A (2021) *On the genesis of thought and language*. Academic Studies Press, Boston
- Kugurakova V, Talanov M, Manakhov N, Ivanov D (2015) Anthropomorphic artificial social agent with simulated emotions and its implementation. *Procedia Compu Sci* 71:112–118. <https://doi.org/10.1016/j.procs.2015.12.217>
- Letheren, K., Russell-Bennett, R., & Whittaker, L. (2020). Black, white or grey magic? Our future with artificial intelligence. *Journal of Marketing Management*, 36(3–4), 216–232. <https://doi.org/10.1080/0267257x.2019.1706306>
- Lombardo W (2021) *Losing Ourselves*. The New Atlantis. Geraadpleegd op 15 April 2022, van <https://www.thenewatlantis.com/publications/losing-ourselves>
- Maj M (2007) Are we able to differentiate between true mental disorders and homeostatic reactions to adverse life events? *Psychother Psychosomatics* 76(5):257–259
- Malinowska JK (2021) What does it mean to empathise with a Robot? *Mind Mach* 31(3):361–376. <https://doi.org/10.1007/s11023-021-09558-7>
- Massumi B (1995) The autonomy of affect. *Cult Crit* 31:83–109
- Misselhorn C (2009) Empathy with inanimate objects and the Uncanny Valley. *Mind Mach* 19(3):345–359. <https://doi.org/10.1007/s11023-009-9158-2>
- Mori M (1970) Bukimi no tani, *Energy* 7(4), pp. 33–35, translated into English by K.F. MacDorman and T. Minato (2005). *Proceedings of the Humanoids-2005 workshop: Views of the Uncanny Valley*, Tsukuba, Japan
- Mori M (2005) On the uncanny valley. *Proceedings of the Humanoids-2005 workshop: Views of the uncanny Valley*, Tsukuba, Japan.
- Mower E, Metallinou A, Lee CC, Kazemzadeh A, Busso C, Lee S, Narayanan S (2009) Interpreting ambiguous emotional expressions. 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops. <https://doi.org/10.1109/acii.2009.5349500>
- Ott BL (2017) Affect in Critical Studies. *Oxford Research Encyclopedia of Communication*. <https://doi.org/10.1093/acrefore/9780190228613.013.56>
- Parviainen J, Coeckelbergh M (2020) The political choreography of the Sophia robot: beyond robot rights and citizenship to political performances for the social robotics market. *AI & Soc* 36(3):715–724. <https://doi.org/10.1007/s00146-020-01104-w>
- Pelau C, Ene I, Pop MI (2021) The impact of Artificial Intelligence on consumers’ identity and human skills. *Amfiteatru Economic*. 23(56):33. <https://doi.org/10.24818/ea/2021/56/33>
- Quiroga-Clare C (2003) Language Ambiguity. *Translation Journal*. Geraadpleegd op 15 april 2022, van <https://translationjournal.net/journal/23ambiguity.htm>
- Ravetto-Biagioli K (2016) The digital uncanny and ghost effects. *Screen* 57(1):1–20. <https://doi.org/10.1093/screen/hjw002>
- Ravetto-Biagioli K (2019) *Digital Uncanny*. Oxford University Press, Oxford
- Salvini P (2015) Of robots and simulacra: the dark side of social robots. In I. Management Association (Eds.), *Human rights and ethics: concepts, methodologies, tools, and applications* (pp. 1424–1434). IGI Global. <https://doi.org/10.4018/978-1-4666-6433-3.ch078>
- Samani, H. A., & Saadatian, E. (2012). A Multidisciplinary Artificial Intelligence Model of an Affective Robot. *International Journal of Advanced Robotic Systems*, 9(1), 6. <https://doi.org/10.5772/45662>
- Spatola N, Wudarczyk OA (2021) Ascribing emotions to robots: explicit and implicit attribution of emotions and perceived robot anthropomorphism. *Comput Hum Behav* 124:106934. <https://doi.org/10.1016/j.chb.2021.106934>
- Thomason TC (2014) Criticisms, benefits, and limitations of the DSM-5. *Arizona Counseling Journal*, 30. http://works.bepress.com/timothy_thomason
- Turkle S (2018) *The Assault on Empathy*. Behavioral Scientist. Geraadpleegd op 15 April 2022, van <https://behavioralscientist.org/the-assault-on-empathy/>
- Turkle S (2021) *The Empathy Diaries: A Memoir*. Penguin Books
- Vermeulen P (2014) Posthuman affect. *Eur J Engl Stud* 18(2):121–134. <https://doi.org/10.1080/13825577.2014.917001>
- Wieseke J, Geigenmüller A, Kraus F (2012) On the role of empathy in customer-employee interactions. *J Serv Res* 15(3):316–331. <https://doi.org/10.1177/1094670512439743>
- Zhou, Z. (2021). Emotional thinking as the foundation of consciousness in artificial intelligence. *Cultures of Science*, 4(3), 112–123. <https://doi.org/10.1177/20966083211052651>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.