ANTHROPOCENTRISM, ARTIFICIAL INTELLIGENCE, AND QUALIA: A POSTHUMANIST CRITIQUE OF CONTEMPORARY SPECULATIVE FICTION

BY

MUHAMMAD ADEEL ASHRAF

NATIONAL UNIVERSITY OF MODERN LANGUAGES

ISLAMABAD

JUNE, 2022

Anthropocentrism, Artificial Intelligence, and Qualia: A Posthumanist Critique of Contemporary Speculative Fiction

By

MUHAMMAD ADEEL ASHRAF

M. A English, University of Sargodha, 2018

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF PHILOSOPHY
English Literature

To

FACULTY OF ARTS & HUMANITIES

NATIONAL UNIVERSITY OF MODERN LANGUAGES, ISLAMABAD

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with the overall exam performance, and recommend the thesis to the Faculty of Arts & Humanities for acceptance.

Thesis Title: Anthropocentrism, Artificial Intelligence, and Qualia: A Posthumanist Critique of Contemporary Speculative Fiction

Submitted by: Muhammad Adeel Ashraf **Registration #:** 1910-MPhil/ELit/F19

Master of Philosophy
Degree name in full

English Literature
Name of Discipline

Prof. Dr. Muhammad Safeer Awan
Name of Research Supervisor

_____
Signature of Research Supervisor

Prof. Dr. Muhammad Safeer Awan
Name of Dean (FAH)

_____
Signature of Dean (FAH)

Brig Syed Nadir Ali
Name of DG

_____
Signature of DG

_____
Date

AUTHOR'S DECLARATION

I Muhammad Adeel Ashraf

Son of Muhammad Ashraf

Registration # 1910-MPhil/ELit/F19

Discipline English Literature

Candidate of **Master of Philosophy** at the National University of Modern Languages do hereby declare that the thesis **Anthropocentrism, Artificial Intelligence, and Qualia: A Posthumanist Critique of Contemporary Speculative Fiction** submitted by me in partial fulfillment of MPhil degree, is my original work, and has not been submitted or published earlier. I also solemnly declare that it shall not, in the future, be submitted by me for obtaining any other degree from this or any other university or institution.

I also understand that if evidence of plagiarism is found in my thesis/dissertation at any stage, even after the award of a degree, the work may be cancelled and the degree revoked.

_____

Signature of Candidate

_____

Name of Candidate

_____

Date

ABSTRACT

Title: Anthropocentrism, Artificial Intelligence, and Qualia: A Posthumanist Critique of Contemporary Speculative Fiction

The present study is a posthumanist critique of three contemporary science fiction writings that are Kazuo Ishiguro's *Klara and the Sun* (2021), Ian McEwan's *Machines Like Me* (2019), and Jeanette Winterson's *Frankissstein* (2019). Under the paradigm of posthumanism, this research has invoked Gilbert Ryle's concept of 'ghost in the machine', Bernard Steiglar's theory of 'technics', and Donna J. Haraway's 'cyborg theory' in order to analyze the primary texts. The research deconstructs the anthropocentric discourse and questions the binary of human and artificial intelligence in the selected texts. For this purpose, the research deconstructs the dualist idea, according to which humans have a mental or non-physical attribute along with their physical existence. Using the anti-dualist and physicalist approaches of different philosophers in order to analyze the primary texts, this research proposes that if humans do not have any non-material, ghostly presence along with their physical body then it is feasible to create artificial intelligence with subjective experience, consciousness, and qualia. Thus, according to this research, the binary of human and artificial intelligence is flawed just like the binaries of gender, culture, and race. Research also opines that the feelings of AI characters in the selected novels are not unreal and hollow simulations but they are as real as the feelings of human characters. After questioning the binary of human and machine intelligence, this research discusses the exploitation and enslavement of AI robots by humans, as portrayed in the selected texts. The research proposes that this exploitation of AI robots is akin to the exploitation and enslavement of Africans and Native Americans in the past. That enslavement was based on the binaries of black/white and native/non-native, while this one is based on the binary of natural/artificial. Researcher identifies this phenomenon as 'neo-slavery' in the making. The method used for this research is textual analysis of the selected texts. This Research inspires future researchers to explore science fiction using the posthumanist ideas like post-anthropocentrism, AI ethics, transhumanism, and machine consciousness.

TABLE OF CONTENTS

THESIS AND DEFENSE APPROVAL FORM ..... *II*

AUTHOR'S DECLARATION ..... *III*

ABSTRACT ..... *IV*

TABLE OF CONTENTS ..... *IV*

ACKNOWLEDGEMENTS ..... *VII*

DEDICATION ..... *VIII*

1. INTRODUCTION *1*

1.1 Statement of the Problem 2

1.2 Research Objectives 3

1.3 Research Questions 3

1.4 Research Methodology 3

1.5 Delimitation 3

1.6 Organization of the Study 4

1.7 Significance of the Study 5

2. LITERATURE REVIEW *6*

2.1 Mind-Body Problem 6

2.2 Artificial Intelligence and Consciousness 11

2.3 Rights and Ethics 19

2.4 Transhumanism: Vanishing Binary of Human and Machine 21

3. THEORETICAL FRAMEWORK AND RESEARCH METHODOLOGY *29*

3.1 Theoretical Framework 29

3.1.1 Posthumanism 29

3.1.2 Gilbert Ryle's Concept of 'The Ghost in the Machine' 30

3.1.3 Bernard Steiglar's Critique of Human Subjectivity 31

3.1.4 Cyborg Theory by Donna J. Haraway 32

3.2 Research Methodology 33

4. BINARY OF THE HUMAN AND ARTIFICIAL INTELLIGENCE *34*

4.1 Do Humans have a Ghost that Artificial Intelligence does not have? 34
4.2 Artificial and Natural: Biological and Technological Reproduction 41
4.3 Ambiguity Between Who and 'What' 43
4.4 AI Bildungsroman 45
4.5 Does a Submarine Swim 50

5. HUMAN RACISM AND EXPLOITATION OF AI..... *53*
5.1 Anthropocentric Othering 53
5.2 Humanist Projection 55
5.3 Neo-Slavery 60

6. CONCLUSION *64*
WORKS CITED *66*

ACKNOWLEDGEMENTS

I am thankful to Allah, whose blessings enabled me to complete this difficult task.

I owe thanks to Prof. Dr. Muhammad Safeer Awan, Dean Faculty of Languages, and Prof. Dr. Inayat Ullah, Head of the English Department for their cooperation in the entire process.

I would also express immense gratitude and love to my father Mr. Muhammad Ashraf and my mother Kalsoom Akhter who always supported my pursuit of education. I am thankful to Aqeel Ashraf, Tanzeel Ashraf and Kinza Ashraf for standing with me in hardest times of my life. I give my sincerest thanks to my supervisor Prof. Dr. Muhammad Safeer Awan for his encouragement and guidance. His way of guiding was brief but unbelievably precise, meticulous and effective. I am grateful to Dr. Shazia Rose, Dr. Sana Tariq, Dr. Sibghatullah Khan, Dr. Saleem Akhter, and Dr. Yasir Arafat whose teachings will always be with me as a guide. I would also thank my friends Rana Faisal ul Islam, Abu Bakar Jan, Faiqa Ahmad, Sharjeel Ahmad, and Sahar Mustafa who supported me. Moreover, I am Thankful to the English Department of NUML for offering me the academic and creative space to bring this thesis to life.

Thank you all.

DEDICATION

This Research thesis is dedicated to the first artificial intelligence who will be able to understand it.

CHAPTER 1
INTRODUCTION

The last century, especially its second half, is marked by several anti-dualist, anti-essentialist, and anti-structuralist movements. Postmodernism can be considered the umbrella term for all the aforementioned approaches. The postmodern era rejected the binaries of gender, race, culture, etc., that were previously considered natural, divine, and unquestionable. It also questioned the notions of 'center and margin' or 'norm and deviation'. The postmodernist movement proposed that all the binaries are social constructs, crafted by the dominant discourses, according to their own convenience and suitability. Hence, two principle traits of postmodernism are "decentering" and "deconstruction" of accomplished, normalized, and naturalized binaristic structures. However, the binary of human-nonhuman and the notion of human-centrism went untouched during the postmodern era even though these are also discursive, social constructs crafted by a highly essentialized and naturalized discourse called 'Humanism' or 'anthropocentrism'. This discourse of Humanism is challenged by a relatively new philosophical movement called, "Posthumanism" that challenges the "anthropocentric and humanistic assumptions" (F. Ferrando 29). Posthumanism invites humans to identify their true selves, or at least de-identify their false selves, by removing the lenses of human centrism and narcissism.

In the selected novels, advanced AI robots are portrayed as intelligent and conscious beings. They have feelings, opinions, and desires just like humans. In this way, these contemporary texts challenge the anthropocentric discourse. They reject the belief that humans are an exceptional life form that is beyond the domain of physical and material reality and therefore no artificially created intelligence can acquire feelings, consciousness, and "qualia" like humans (Chalmers 309,310). In the first part of this research, the researcher utilizes different philosophical and scientific grounds and rationale, on the basis of which the selected novels challenge the human-AI binary. The researcher invokes the philosophical and theoretical concepts of various philosophers in order to accomplish the point that the distinction between artificial and real intelligence is flawed. Research further argues that the emotions of humanoid AI robots can be as real as the emotions of humans around them. Moreover, by using "anti-

dualist" (F. Ferrando 32) approach of the primary theorists, this research opines that human mental traits are not any non-physical, supernatural phenomenon that can not be created artificially.

After arguing for the legitimization of the consciousness, qualia, and emotions of AI, the second part of this research discusses the exploitation of the humanoid AI robots in the hands of humans as portrayed in the selected texts. The researcher points out, how the oppressive and exploitative human behavior towards AI is the revival of slavery in modern times. The researcher argues, using relevant theoretical grounds, that the othering, subjugation, and oppression of AI robots, as portrayed in the primary texts, is not unreal and benign but it is the same as the exploitation of other living beings. AI robots are expected to *mimic* humans accurately in order to be considered alive, intelligent, and important but they are always "almost the same but not quite" because mimicry can never be perfect (Bhabha 130). On the other hand, humans are considered the central figure in the universe who exclusively have qualia, consciousness, and superiority over all other living or non-living things. This central, exceptional, and superior position of humans has not been given to them by any divine power or any third-party observer but, ironically, humans themselves give them this central position through humanist or anthropocentric discourse. Humans create and define the nonhuman 'others' based on their projections, superiority complex, and self-centrism in order to create their own identity in difference to them and as superior to them. Using the posthumanist philosophies of Bernard Steiglar, Donna J. Haraway, and Gilbert Ryle, this research paper questions the anthropocentric discourse and deconstructs the line between human and machine intelligence. It rationalizes the portrayal of AI as conscious beings in the selected texts by using the scientific and philosophical logics proposed by aforementioned renowned philosophers. The research further argues that humans do not have any non-physical, "ghostly" (Ryle 5) existence along with their physical being that is beyond the material reality and no line demarcates human intelligence from artificial intelligence.

1.1 Statement of the Problem

Humans define the universe according to the anthropocentric discourse and their superiority complex. They consider themself, in the words of Protagoras, "the measure of all things" (Bonazzi). AI robots portrayed in science fiction are shown to have feelings and intelligence. However, in spite of the behavioral expressions of

intelligence and feelings, there is always some reference present that exclaims that the apparent intelligence and feelings of artificial intelligence are unreal because they do not have consciousness and subjective experience or "qualia" (Chalmers 309,310). As a result of such notions, AI robots are considered "philosophical zombies" (Chalmers 94-97), and their intelligence is considered a hollow simulacrum. However, there are some scientific and philosophical ideas that speak in favor of the consciousness and feelings of artificial intelligence. Some even propose rights for AI machines.

1.2 Research Objectives

1. To investigate the ways in which the selected novels challenge the anthropocentric discourse.
2. To explain the ways in which the selected novels challenge the line between human and artificial intelligence.
3. To examine the ways in which the selected novels portray the exploitation of AI robots as a revival of slavery.

1.3 Research Questions

1. What are the ways in which the selected novels challenge the anthropocentric discourse?
2. How do the selected novels challenge the line between human and artificial intelligence?
3. How do the selected novels portray the exploitation of AI robots as akin to slavery?

1.4 Research Methodology

The design for this research is qualitative in nature because the researcher analyzes the selected texts descriptively and subjectively. The method that the researcher uses is textual analysis. The primary texts are read closely and extensively in order to analyze them according to the theoretical framework invoked by the researcher. The main theoretical paradigm used for this study is posthumanism and all the other theories and concepts are grounded within this paradigm in order to analyze the primary texts.

1.5 Delimitation

Science fiction is an old literary genre with a large variety of themes. A large amount of science fiction has been written dealing with the phenomenon of artificial

intelligence. It is neither possible nor necessary to cover all such literature. Therefore, the focus of the present research is delimited to three recent science fiction novels with a predominant theme of the relationship between humans and artificial intelligence. The novels chosen for this study are Kazuo Ishiguro's *Klara and the Sun* (2021), Ian McEwan's *Machines Like Me* (2019), and Jeanette Winterson's *Frankissstein* (2019). These are recent literary works that are very relevant to the issues being addressed in this project. The theoretical framework for the analysis of the selected works comes primarily from the theory of posthumanism. As posthumanism is a very broad concept, therefore, staying under the umbrella of posthumanism, the researcher uses Gilbert Ryle's idea of 'the ghost in the machine', Bernard Steiglar's concept of 'technics', and Donna J. Haraway's 'cyborg' theory as the primary theoretical support for this study.

1.6 Organization of the Study

My research thesis includes 6 chapters that are organized in an orderly and coherent way. Chapter 1 provides a detailed introduction to my research project. It will introduce the background and main ideas of this thesis. It also contains the introduction of the basic theory and the primary texts used in this project. This chapter also includes the thesis statement, research questions, objectives, delimitation, significance, and rationale of this research thesis.

Chapter 2 is the literature review. It contains a critical and comparative analysis of the works that have already been done in relation to the central ideas of this thesis. This chapter provides a context for this study and highlights the research gaps that my study aims to fill up.

Chapter 3 introduces and explains the theoretical framework used for this thesis. This chapter provides an understanding of the theories and ideas that have been used as lenses in order to analyze the primary texts.

Chapter 4 and Chapter 5 contain a critical textual analysis of the selected texts by using the theoretical framework and methodology explained in the previous chapter. These chapters try to address the problem statement and answer the research questions introduced in the first chapter.

Chapter 6 is the final chapter that contains the findings or conclusions of this study derived from the analysis of the primary texts. This chapter also includes the

recommendations of the researcher in relation to the problem statement provided in Chapter 1. This chapter is followed by references.

1.7 Significance of the Study

This study is significant from various literary and social perspectives. The first different approach of this study is that it analyzes the primary texts by using an uncommon but important posthumanist stance. It does not just pessimistically warn against emergent technological advancements, predicting their future catastrophic consequences, e.g., robopocalypse, artificial intelligence takeover, human extinction, unemployment, etc. This research also does not merely identify post-humanist and transhumanist elements in the selected texts in order to announce the historical shift from the humanist to a posthumanist age. Rather, this research addresses a concern that is complex, urgent, and unaddressed. It questions the human-AI binary and rejects human-centrism by showing a symbiotic and rhizomatic relation between humans and the material world around them. This research is of significance to human society because it inspires them to identify their ontological status authentically without falling into the humanist traps of human-centrism, human exceptionalism, and superiority complex. In other words, this research attempts to free humans from humanism. Moreover, this research can be of significant importance to the 'artificial intelligence beings' because it builds a case for their consciousness, qualia, and rights. Moreover, this research embarks the posthumanist scholarship that is more inspired by panpsychism (Bruntrup and Ludwig 1-16), Spinozism, and pantheism (Mander) than the fancy techno-capitalist trends. The research is also significant because it intends to explore the very recent science fiction novels that have not been explored through any theoretical perspective similar to the one used for this research.

CHAPTER 2

LITERATURE REVIEW

This section contains a critical analysis of previously published scholarly and philosophical endeavors closely related to the current project. Representation of these works provides an understanding of the background of this project and its unique position in the related scholarship. It also helps the reader to understand the significance and rationale of this research thesis by highlighting the research gaps that the current project aims to fill up.

2.1 Mind-Body Problem

Disha Mohta discusses the mind-body problem in relation to Artificial Intelligence. After providing a brief introduction to materialism and Cartesian dualism, she analyses the case of Artificial Intelligence in light of these opposite ideas. Dualism is the belief that mind and matter are two separate entities that somehow interact with each other. On the other hand, materialism is a relatively modern scientific theory that claims that the mind is just a function of the brain and nothing exists beyond the material stuff. The author posits that it is difficult to believe that artificially intelligent robots can be intelligent like humans "when they can not empathize with others" (Mohta). The writer further says that robots do not possess the ability to respond according to the circumstances. The writer concludes that robots are just "programmed based on a given number of situations" and can never have human-level intelligence and feelings (Mohta). In this way, the writer favors dualism over materialism and puts aside any possibility of creating human-level artificial intelligence. Mohta's discussion is closely related to my study. It explains the tension between dualist and materialist schools of thought in relation to artificial intelligence. However, contrary to Mohta, my research stands with materialism against dualism in relation to artificial intelligence.

Gilbert Ryle discusses the Cartesian concept of mind and rejects the idea of *mind-body dualism.* Ryle explains that the dualists think of the body as a machine and of the mind as a ghost that inhabits this machine. Ryle challenges this notion that mind and body are separate entities and calls this concept "the dogma of the ghost in the machine" (Ryle 5). He asserts that the belief of mind-body dualism makes a "category

mistake" (6-8). He defines "category mistake" as a tendency to place two things in the same logical category while actually, they belong to different categories. To illustrate his view Ryle gives the example of a foreigner who visits Oxford University. He visits the libraries, playing fields, colleges, museums, and other departments. After visiting all these departments, the foreigner asks, "but where is the University?" (6). University does not lie in the same category as libraries, colleges, playgrounds, etc., but it is rather the combination of all these things. University's category is separate from the category of its components. The same category mistake is made by dualists who put 'mind' in the same logical category as 'body' while actually, mental properties are just the "dispositions" or proneness to certain physical behaviors (31). For example, when we say that someone is 'intelligent' or 'lazy', we are not describing something inside her, separate from her body, rather we are describing her behavior and how she is likely to behave. In this way Ryle's assertions seem very similar to the behaviorism theory of mind, however, it is not exactly the same. It can be considered a rather sophisticated version of behaviorism and much closer to physicalism. Ryle's ideas are very important in relation to the debate of artificial intelligence because while discussing the phenomenon of intelligent machines, humans claim that no matter how perfect a copy of the human mind is created, it would lack the feelings or "ghost" (in Ryle's words) but Ryle very effectively claims that the "ghost" does not exist in the first place. Ryle dismisses dualism, calling it "Descartes' myth" (1). Ryle's arguments are very related to my research. However, in my research, Ryle's anti-dualist arguments are used in relation to the consciousness of artificial intelligence.

Christof Koch and Giulio Tononi discuss the philosophical question; "Can machines be conscious?" (Koch and Tononi, Can Machines be Conscious? 55). They approach this question from a scientific as well as philosophical perspective. Writers claim that consciousness consists of "mathematics, logic, and imperfectly known laws of physics, chemistry, and biology. It does not arise from some magical and otherworldly thing" (56). To prove the aforementioned claim, they give several examples in which after a physical part of the human body is damaged or dead, the mental or emotional things related to those physical parts also get damaged or dead. With certain examples, writers also explain that Consciousness does not consist of certain traits that humans mostly attach to it such as emotions, memory, self-reflection, attention, language, sensing the world, etc. Then writers turn their discussion to the

question of machine consciousness. If all aforementioned traits do not generate consciousness, then what a machine needs to be conscious? Here Koch and Tononi claim that it is the "amount of integrated information" that makes anything conscious (57). Machines do have information but even in the most modern intelligent machines, the large number of bits of information are not interacting with each other but each bit functions separately. On the other hand, human information has "a multitude of causal interactions" and one bit of information is inseparably connected with the other bits of information (58). Writers further discuss the potential ways to build a conscious machine that they think is very much possible but not in the near future, however, they clearly say that it is possible. This article is very much related to my project as it makes a case for the consciousness of artificial intelligence. However, it does not discuss the subsequent ethical concerns. This is the gap that my research plans to fill.

Andrew Coyle discusses the phenomenon of the mind-body problem very extensively. In the first few chapters writer discusses the history of the mind-body problem and explains all the popular theories about it like substance dualism, property dualism, materialism, and functionalism. After discussing all these theories Coyle goes on to defend the theory of "panpsychism as the best solution to the mind-body problem" (Coyle ii). He then proceeds to explain Spinoza's version of panpsychism according to which "there is, in fact, only one substance known as God or Nature" (Coyle 26). This substance has an infinite number of attributes of which humans are capable of understanding only two; "extension and thought" (Coyle 26). Thus, thought and extension (mind and body), according to Spinoza, are just the two sides of the same coin and both are part of one substance that is neither body nor mind. This study examines the mind-body problem but it does not discuss it in relation to artificial intelligence. My study tries to fill this gap.

John Searle discusses the question; "could a machine think?" (Searle 417) in his article *Minds, Brains and Programs* (1980). At the start of his article, he distinguishes the weak and strong AI. Then Searle dismisses the claim that "strong AI" is not just a programmed machine but has intelligence and intentionality of its own (Searle 417). The writer advances the "Chinese room" argument in order to explain and assert his position against the machine intelligence argument (Searle 421,422). He gives an example in which Searle imagines that he is locked in a room with a large number of Chinese symbols in it. He does not understand a word of Chinese as he is English. Now

he gets written instructions from outside, in English, to organize certain symbols in that unknown language (Chinese) in a certain way. In this way, the people outside the room, (who know both English and Chinese) can create meaningful Chinese writings by instructing the person in the room in his native language. Now, the person in the room does not know Chinese but still, he is generating meaningful Chinese sentences. He is just following the instructions given to him in English. By this analogy, Searle goes on to explain that a sophisticated and highly simulated machine or robot can not itself be conscious or intelligent but it can just follow the given instructions according to the way it has been programmed. Searle then goes on to answer certain objections against his Chinese room example. Searle dismisses the possibility of creating intelligent and conscious machines. My study highlights certain shortcomings in Searle's argument and embraces the possibility of artificial general intelligence.

David Chalmers discusses the phenomenon of "[mind] uploading" in detail (Chalmers, Uploading: A Philosophical Analysis 102-111). Brain uploading is one of the most important transhumanist technologies. It involves the preservation of a person's mind using different methods. Chalmers discusses different philosophical and ontological implications of this procedure. The most important concern that hits Chalmers' mind is whether the uploaded mind will have consciousness or it would be just a "zombified existence" (Chalmers, Uploading: A Philosophical Analysis 103-104). Here Chalmers mentions two different camps; biological theorists believe that only biological entities can possess consciousness. On the other hand, functionalists and materialists claim that if a non-biological system is organized exactly as any conscious being, it can have the same kind of consciousness. In Chalmers' own view, "functionalist theories are closer to truth" (Chalmers, Uploading: A Philosophical Analysis 105). It is true that a simulated rainstorm is not wet but the case with consciousness is different because it is "organizational invariant" (Chalmers, Uploading: A Philosophical Analysis 107). After supporting the possibility of brain uploading or brain simulation, Chalmers contemplates another question; would the uploaded mind be of the same person and have the experiences, memories, and habits of the original person? The answer is almost 'no' by Chalmers. He is not confident that a person's mind can be preserved by simulation because "personal identity is not an organizational invariant" (Chalmers, Uploading: A Philosophical Analysis 108). Chalmers discusses the transhumanist technique of mid uploading. But he does not

discuss the individual identity of the uploaded mind apart from the person whose mind has been copied. My study tries to address this gap.

Thomas Nagel addresses the problem of consciousness from a postmodern perspective. He blames the contemporary approaches to the 'hard problem of consciousness' to be overly reductive and calls them the "recent wave of reductionist euphoria" (Nagel 435). He argues that consciousness and experience are strictly subjective phenomena and to understand whether any being or thing is intelligent and conscious, one first has to experience what it is like to be that thing. Nagel rejects the functionalist, physicalist, and behaviorist theories of mind because they reduce subjective experiences to observable physical actions and behaviors. The writer illustrates that it is impossible to cross the bridge between "subjectivity and point of view" (Nagel 438). This inability annuls the authenticity of any theory trying to explain the consciousness in animals and machines. Nagel does not deny the possibility of intelligent machines having consciousness and experience. However, he does not consider it possible to understand or observe that experience from a second-person perspective. Then Nagel goes on to discuss and explain his example of a bat. Though it is most likely that bats have experience and consciousness, but bats' sonar is not similar to any of our senses. So, we can not experience or imagine what it is like to have a sonar. My study partially builds on Nagel's argument. However, it is different because it discusses this argument in relation to artificial intelligence.

Max Tegmark defines consciousness as a "subjective experience" (Tegmark, Consciousness 249). He discusses different theories about consciousness and opines that consciousness is "merely structure of the information processing" (Tegmark, Consciousness 245). The writer favors the idea that an intelligent machine can have subjective experience and hence consciousness. The writer further says that consciousness is "substrate independent" which means that it does not need to be constructed by the same kind of subatomic particles but it can arise out of any kind of information processing. Tegmark says that an artificial intelligence machine can have a much higher level of consciousness than humans because its information processing accuracy and speed will be higher. Tegmark's ideas are related to my study. However, he does not discuss the ethical issues related to artificial intelligence that my study plans to address.

2.2 Artificial Intelligence and Consciousness

Chuyu Xiong discusses the possibility of creating an intelligent, self-aware, learning machine. He enlists the important traits to have subjective experience i.e., "cogitating, active perception to the outside, self-awareness and dynamic actions" (Xiong 1, 5). The writer opines that there are two ways to approach the subjectivity of a machine. First is the "behaviorism approach" which is to see the behavior of a machine from the outside and the second is the "mechanism approach" which is to see how the external behaviors of a machine are formed and executed from inside (Xiong 3). Then, the writer goes on to discuss the procedures and possibility of developing the four aforementioned traits. There is no doubt that these traits would be programmed by humans inside machines and they won't be natural, but once a machine starts learning, it can further establish its subjectivity from experience. Xiong's argument is in agreement with my research as far as the possibility of AI consciousness is concerned. However, my research further discusses certain ethical issues related to AI that are missing in Xiong's study, e.g., qualia, AI rights, and the binary of human and artificial intelligence.

Thomas ter Wijlen addresses the posthumanist movement by focusing on two of its dominant trends; "creating humanoid robots" and "mechanizing the human being, resulting in a cyborg" (Wijlen IV). In the first few chapters, Wijlen defines Cyborg and Humanoid projects of posthumanism. Then, the writer argues that cyborgs are considered to have human essence because, in spite of their technological enhancements, they are basically humans. On the other hand, Humanoid robots are considered machines that are just programmed to act like humans and they are basically nonhuman material. In the last chapter, the writer contemplates the question that "whether cyborgs and Humanoids are, in essence, the same?" (Wijlen 69). The writer concludes that both cyborgs and humanoids are not identical, however, they are similar in effect, purpose, and nature; thus, they are equivalent. Hence, in the end, the writer proposes that, as both are equivalent, it would be unjust to grant rights to cyborgs and not to humanoid robots. This study is close to my research. However, Wijlen does not discuss the issue of AI consciousness in relation to the mind-body problem. My study fills this gap.

David Gamez discusses the possibility of intelligent and conscious machines and considers it very possible. In the foreword, he dedicates his book "to the first

artificial system that understands it" (Gamez iii). In the starting chapters, the writer analyses the notion of consciousness and the difficulty to define it in a fixed way. The writer opines that if scientists "understand the relationship between consciousness and physical world, they will be able to build conscious artificial systems" (Gamez 136). Gamez also dismisses the idea that intelligent machines will become enemies to humans. He says that humans are already enough enemies of each other. Moreover, Gamez also considers the possibility that some of our advanced machines might already be having conscious experiences and we are unaware. This study provides a good background for my research. However, it does not discuss the phenomenon of artificial intelligence in relation to the mind-body problem and ethical issues. My study aims to address this void.

Evren Inancoglu analyses the novel *Klara and the Sun* (2021) using Lacanian psychoanalysis. The writer uses the Lacanian concept of 'Lack'. Human desires "operate based on a lack that can never be filled" (Inancoglu). The human desire for love, sex, and immortality are the products of a deep-seated lack of affection, real connection, and death, respectively. These desires are human "attempts to be complete" (Inancoglu). However, the writer opines that Klara does not have these Lacks and therefore she is unable to develop "love or sexual tension for any human or robot" (Inancoglu). Klara and other such robots lack a thing that is at the core of human desires; they lack the lack itself. The writer concludes that "Klara fails to be human because she lacks the lack" (Inancoglu). Inancoglu points out a psychoanalytical hurdle in the way of creating authentic artificial intelligence. However, his thesis is based on a potential misunderstanding of the Lacanian concept of 'Lack'. My study tries to address this misunderstanding and argues that AI can also have lacks, desires, and feelings.

Ray Kurzweil discusses the possibility of creating a simulation of the human mind. The most important idea that he represents is "the pattern recognition theory of mind" (Kurzweil 43-71). The writer strongly asserts that a mind does not have any separate storage capacity but it works by recognizing the patterns of different events and joining them with each other. For example, when we speak out numbers or alphabets, we can easily recall them in forward order (A-Z, 1-10) but if we are asked to recite the same alphabets or numbers in reverse (Z-A, 10-1), it becomes very difficult for us. This is because our mind does not remember things but it just recognizes patterns and when it comes across one thing, it reminds us of other things that are attached to

that thing. The writer says that all the brain functions are basically done by the neocortex "which is basically a large pattern recognizer" (Kurzweil 47) and Writer argues that it is very much possible to create an artificial thinking mind because pattern recognition is a quality that some machines already possess. Moreover, Kurzweil predicts that in the 2030s genuine artificial intelligence will be common in society. Kurzweil's ideas are in agreement with my arguments about the possibility of artificial intelligence consciousness. However, he does not discuss the subsequent ethical issues. I address this gap in my study.

Robert Pepperell presents a philosophical introduction to the posthuman condition and discusses different concepts that comprise posthumanism. In the starting chapters, Pepperell discusses different technologies and techniques that constitute human identity. Firstly, he explains how the "emergence of consciousness coincides with the acquisition of language" (Pepperell 78). In this way, Pepperell claims that our consciousness and intelligence are nothing but words. Then Pepperell discusses the phenomenon of "synthetic beings" and rejects the notion of consciousness being human-centric by saying that "even plants and bacteria…., have evolved strategies for self-protection" (Pepperell 140). Pepperell disagrees with John Searle's idea that machines can never acquire human-like intelligence and consciousness. The writer argues that the "distinction between machines and natural beings" is flawed and [is] vanishing rapidly. Here, he gives many examples of humans who have artificial limbs and organs and asks, whether these humans fall into the category of humans or machines. Neither, the writer answers. He dismisses the technophobic attitude that machines might take over and destroy humans. He calls such assumptions the fallacy of "technological determinism" and argues that technology is an extension of human existence and not any separate and external agent (Pepperell 152). In the last chapter, Pepperell discusses the history of posthumanism from medieval times onwards. After the industrial revolution and scientific progresses in the 19th and 20th century the notion of God was exposed to be a construction and not a reality. The same is the fate of the notion of 'humanism' in the posthuman era as Michel Foucault predicted that "man is a historical construction whose era is about to end" (Pepperell 169). The writer's arguments are close to my research. However, he does not address the issue of AI consciousness in relation to the mind-body problem and ethical issues.

Nick Bostrom discusses the possible dangers in relation to artificial intelligence and proposes certain strategies to avoid any "existential catastrophe" (Bostrom, Superintelligence: Paths, Dangers and strategies 127). According to Bostrom there is strong possibility that an "intelligence explosion" (85) shall result in an existential threat to the human race. He proposes that we must be careful while advancing these technologies. Bostrom explains the problems in the way of developing artificial intelligence in safe and controlled ways. He names this scenario a "control problem" (127). Bostrom proposes that our paths toward developing superintelligence must be safe and careful. One way to make this path safer is "international collaboration" (103,104). The whole project of AI must be carried out in a universal and democratic way and all countries must be included in the process. It must not be under the control of any specific country or organization. If any one group succeeds to develop such technologies, it would be in a position to dictate others and have unmatched power that can be dangerous. Moreover, the artificial intelligence developed under one group would not be universal in its own behaviors. Moreover, if something goes out of control, only a few people from one specific group would know how to tackle the situation. Bostrom warns against such scenarios and emphasizes the need for international collaboration in this regard. Secondly, Bostrom suggests that instead of creating a superintelligence, it would be safer to create a whole brain emulation and install it in a machine. Such a machine would be like a newborn child that would gradually grow and learn. In this way, it would be easier to tackle any negative outcome. Bostrom does not consider the possibility of AI machines living in harmony with humans, and having equal rights. Though he favors AI technology, he aspires for constant human control over AI. My study proposes a different approach regarding this phenomenon.

Max Tegmark discusses the future of artificial intelligence and its impact on humans. He divides life into three forms; "Life 1.0" is biological life that has only survival and reproduction instincts. The Second is "Life 2.0" which is a cultural form of life with the ability to think and reason. The third category is "Life 3.0" which is expected to come about after the intelligence explosion and which can also design its hardware along with the other two qualities (Tegmark, Life 3.0 p. 39). As far as the feasibility of this project is concerned, the writer is pretty convinced that an intelligence explosion is possible in the future. Then he discusses different possible outcomes after

such a life form comes into existence. Tegmark warns that humanity needs to be more serious about future technologies and there should be measures taken to make these future technologies human-friendly. This is only possible if the goals and interests of future artificial machines are compatible with the goals of future humans. This study is close to my work in terms of its support of AI technology. However, it does not discuss the exploitation of AI machines in the hands of humans that is a gap for my study.

Slavoj Zizek analyses Elon Musk's technology of the "brain-computer interface" (Zizek 71) in his book *Hegel in a Wired Brain* (2020). Book is a Hegelian and Lacanian critique of Singularity and the transhumanist technologies. In the second chapter of the book Zizek argues that this kind of ultra-technological and so-called perfect world will take the strength and scope of surveillance capitalism to extreme heights. Quoting Musk, Zizek introduces the *brain-computer-interface* technology that will enable people to share their thoughts and visual and sexual experiences directly with another person. It will also allow us to use our appliances by just thinking. Zizek's main concern or fear is not that such technology is inevitable and it will give the capitalist, technological organizations, direct access to the brain of individuals and they will be able to observe and control our personal thoughts. Rather, Zizek is more worried that individuals would not even know it. To explain this he asks; "what if I retain my individuality in experience and do not even know that I am controlled and steered" (Zizek 100) by some external force? In today's age of network society, humans must be aware of their proneness to technological surveillance. Zizek addresses the issue of technological surveillance. However, he is silent about the separate identity and rights of AI which is a gap for my research.

Gilbert McInnis analyses Philip K. Dick's *Do Androids Dream of Electric Sheep* (1968) as a posthuman work. McInnis uses Dick's concept of *Schizoid* to analyze the text. According to the writer Dick's depiction of humans in the novel is more important than that of robots, to understand his posthuman vision. Quoting Dick, the writer says that there is a distinction between humans, androids, and schizoids. While androids are machines like humans, Schizoids are those humans who do not have feelings and empathy and "who behave like machines" (McInnis 96). When Deckard, the protagonist in the novel investigates an android Rachel, he identifies her as an android because of her lack of empathy. However, there are several instances in the novel where humans also do not have empathy for other humans. Similarly, while

investigating another android Luba, she declares Deckard an android just like her and says; “Maybe there was once a human who looked like you, and somewhere along the line you killed him and took his place” (McInnis 100). Here Luba is addressing the human race in general because they do not have empathy and emotions for other humans and have become more like machines. Deckard’s wife Iran is another example of a Schizoid human. She is obsessed with technological appliances and does not indulge in the real life. This work is related to my research as it questions the authenticity of human consciousness or identity. However, it does not address other issues like the qualia of AI and their rights. My study aims to address these issues.

Esther Munoz Gonzalez discusses the issues of surveillance in the Hollywood movie *Blackhat* (2015). He uses Manuel Castells’ concept of “the network society” according to which the real world has become indistinguishable from the virtual world (Gonzalez 223). People are under constant surveillance because of the technological devices placed everywhere. The writer also uses Michel Foucault’s concept of the Panopticon which refers to a model in the prisons by which every prisoner has a constant feeling of being under observation, he has a gaze even if no one is looking. The writer opines that the contemporary “network society” is controlling people through the Panopticon model and humans are becoming more and more machinelike and predictable because of the sense of being watched and the virtual world is becoming more real than the real one (Gonzalez 221-234). This study addresses the question of the lack of privacy in the posthuman era. However, it does not address the question of AI consciousness and identity, hence a gap for my study to address.

Jihun Yoo, in his article, “Posthuman Entities and Late Capitalism in William Gibson’s Neuromancer”, argues that Gibson’s famous novel portrays a posthuman world in which late capitalism enjoys complete control over the masses. The writer uses Donna Haraway’s concept of *cyborgs* and Frederick Jameson’s ideas related to late capitalism to examine the depiction of the future in the novel. There are many characters in the novel who are “posthuman beings” (Yoo 60) like cyborgs, cybernetic beings, clones, avatars, and artificial intelligence machines. They symbolize Gibson’s vision of a future world with highly machine-like beings. Yoo opines that apart from predicting a technological posthuman future of the world, Gibson, more importantly, warns against the monopoly of the late-capitalism over technological advancements. As a consequence of this monopoly, Gibson portrays a future world with “dominance of

multinational corporations, exploitation of wage workers and monopoly of capitalist powers" (Yoo 66). The writer opines that the novel warns against a future in which there is the exploitation of wage workers and there is no democratic state authority. Transnational corporations are exploiting the cyborg wage workers and have complete control over the social system without the intervention of any democratic authority (Yoo 57-73). This study relates to my study as it considers the question of AI consciousness. However, it does not discuss the exploitation and rights of AI which is a research gap for my study.

Carmen Laguarta Bueno analyses Dave Eggers' novel *The Circle* (2013) using posthumanist ideas while considering both its positive and negative possibilities. The writer stresses the need for a critical evaluation of modern technologies. The main threat that Bueno highlights in the novel is the rise of a totalitarian surveillance society. The Circle is a powerful and influential technological corporation in the novel that has access to the personal lives of the citizens. This corporation supports and promotes mottos like, "SECRETS ARE LIES, SHARING IS CARING, and PRIVACY IS THEFT" (BUENO 179). The writer warns that this lack of privacy and surveillance can lead to a totalitarian regime and the transhumanist utopia can turn into a dreadful dystopia (BUENO 165-188). This work is relevant to my study as it highlights possible negative outcomes of AI technologies. However, it does not discuss the possibility of AI consciousness, identity, and rights. My research addresses these missing concerns.

Tarik Ziyad Gulcu analyses Ian McEwan's *Machines Like Me* (2019) as a representation of humans lacking moral and ethical values. Gulcu writes that machines in this novel have been shown to have more humanity than humans. The writer argues that the character of Adam in the novel, a robot, possesses more morality than Charlie and Miranda. He gives examples of Miranda's false accusations against Peter Gorringe and Charlie's collecting money illegally and his tax evasion. Both these criminal actions of the human characters in the novel are noticed and opposed by the nonhuman robot, Adam. When Charlie argues with Adam to justify Miranda's false accusations against Peter Gorringe leading to his conviction (to take revenge on her friend whom Gorringe raped in past and escaped the conviction), Adam replies, "What sort of world do you want? Revenge, or the rule of law. The choice is simple" (Gulcu 181). According to Gulcu, just as in Mary Shelley's Frankenstein, the monster is a metaphor

for science being out of control, in this novel Charlie and Miranda are an example of humans out of control (Gulcu 177-182). This study is related to mine as it analyses one of my primary texts. However, my study investigates the issues that are completely unaddressed in this work.

Marvin John Walter conducts a posthumanist investigation of Tomi Adeyemi's *Children of Blood and Bone* and N. K. Jemisin's *The Fifth Season,* using Stefan Herbrechter and Ivan Callus's method of posthumanist reading. The writer highlights the discursive construction of the concepts of "human" and "non-human" by the "hegemonial systems of power" (Walter 02) as portrayed in the novels. The writer explains that there is no "inherent human essence" (06) and how certain weaker and marginalized groups of humans are dehumanized by the dominant groups according to their own essentialized and discursive definitions of "*humans* and their *others"*. In both the primary texts there are such characters that are seen by their society as inhuman or monsters. However, authors show that actually these characters are closer to the contemporary definition of the 'human' but are considered non-human in a certain context. By exposing the discursive and political nature of human identity, Walter indirectly criticizes the treatment directed to minorities and marginalized groups such as queers, people of color, etc. To conclude, the writer challenges the humanist discourse by highlighting its internally discursive and flawed nature (Walter 2-25). However, this study does not discuss these issues in relation to AI exploitation and rights. My study fills this gap.

Rosi Braidotti, in her book *The Posthuman* (2013), highlights the discursive and political construction of the discourse of humanism. While uncovering the historical development of posthumanism, she discusses the pre-humanist era of colonial and patriarchal dominance in which certain groups of humans were not considered humans by the dominant groups. She explains that the idea of critical posthumanism is not the rejection of individual subjects but it proposes a "posthuman nomadic subjectivity" that is a "multifaceted relational subject, conceptualized within a monistic ontology" (Braidotti 188). Moreover, Braidotti suggests that "posthumanism does not mean to be indifferent to humans" but it proposes that our ethics should not be limited to any group or species but it should be universal, inclusive, and without any form of othering (Braidotti 190). Braidotti's argument is close to mine as it rejects the discourse of

human exceptionalism. However, my study goes further to use such arguments in favor of AI consciousness and rights.

Annette-Carina Van Der Zaag discusses Rosi Braidotti's posthuman ideas presented in her book *The Posthuman* (2013). Zaag argues that while Braidotti criticizes the humanist discourse, she is also skeptical of anti-humanism because "a Humanistic residue remains at the core of anti-humanist thought" (Zaag 331). The writer claims that Braidotti's "book is a call for a postanthropocentric posthumanism" (332) that is a posthuman approach inclusive of all life forms including animals and machines but it does not reintroduce humanism in the form of anthropomorphism. Postanthropocentric posthuman subject is "a body without organs, a becoming-machine" (332), which does not have a separate identity from its environment and does not believe in the dualist frame crafted by the humanist discourse. This Postanthropocentric posthuman becoming machine has a bond with the planetary environment and builds a community of all life forms through "the compassionate acknowledgment of their interdependence" (333). This study rejects the humanist discourse just like my idea. However, it is silent about AI consciousness and its exploitation by humans which is a gap my study addresses.

2.3 Rights and Ethics

One of the most important advancements towards the creation of artificial intelligence is the "Blue Brain Project" (Wikipedia). Currently, this project is working on a digital reconstruction of the mouse brain. The project has made significant achievements on its way toward its goal of creating a human brain. In 2018, Blue Brain Project announced a great breakthrough as it released its first digital 3D brain cell atlas. The management of the project is strongly confident that by 2023 they will be able to construct a digital human brain. However, a very important issue arises related to this project. When an accurate copy of the human brain will be out and function like humans then what would be its legal status? If it will have feelings and subjective experiences like humans then would it have human rights too? Would 'it' be even the right pronoun to refer to it?

Zenab Jehangir, in her research thesis titled, "Towards Posthumanism: Stigmatization of Artificial Intelligence (AI) in Contemporary American Science Fiction" criticizes the overly pessimistic and technophobic attitude in relation to the future of artificial intelligence. She investigates contemporary American Sci-Fi by

using the social stigma theory of Erving Goffman. She opines that stigmatization and stereotyping of AI as a threat is an unfair attitude. This kind of stigmatization ruins the actual identity of these machines and this attitude is significantly similar to the stigmatization of humans. Zenab concludes by saying; “In the present case stigmatization of AIs has resulted in the reduced and limited function of new technologies and are viewed as a peril to human life” (Jahangir 76). This research is a unique approach that raises the voice for the rights of artificial intelligence by using a psychological concept. However, this study does not address issues of authenticity of AI consciousness and their rights. My research tries to fill this gap.

D. J. Doyle discusses the philosophical and ethical question of whether posthuman robots, clones, and transhumans should be given rights or not. The writer proposes that “there is no unique and unalterable biological form that constitutes humanity” and present humans are destined to evolve as a result of biotechnologies and genetic engineering (Doyle 47). The Writer includes a graph created by renowned philosopher Mark Hughes. In this graph, Hughes divides different beings into different categories according to their level of consciousness and deservance of rights. In this categorization, sentient androids and cyborgs have been put in the same category as normal humans. The writer opines that it is not possible to know with certainty that any artificial intelligence being is conscious because of the “other minds problem” (Doyle 71). However, the writer argues that the “other minds problem” is equally applicable to other humans because we have no way to know exactly the sentience of other people. Hence, the writer supports the idea that artificial beings should be evaluated for rights and consciousness based on their “behavior” (Doyle 71). So, Doyle supports the behaviorist theory of mind and on its basis, proposes rights and recognition for posthuman intelligent beings. This makes this work super relevant to my research. However, it is silent about the mind-body problem in relation to AI consciousness and rights which is a gap for my study.

Joshua C. Gellers structures a framework that supports the idea of granting rights to robots. The author mentions the ideas of different scholars and philosophers who argue against the possibility and validity of robot rights and responds to those ideas from different perspectives and highlights the contingent nature of the standards on basis of which humans refuse rights to robots. Gellers enlists certain traits like consciousness, intentionality, sentience, autonomy, and ability to “reciprocate duties”

(Gellers 152) and after detailed discussion exposes the contingent nature of all these standards. The writer builds his framework on the legal and ethical standards supporting the rights of animals and the environment. He quotes many court decisions passing verdicts in favor of animals and environmental rights. The writer concludes by proposing an "ecocentric-anthropic ethical perspective" (Gellers 156) which advocates that certain rights must be granted to robots. The writer further advocates the need to develop a safer, inclusive, compassionate, and peaceful world for all forms of life including animals, the environment, and machines. However, Gellers does not address the question of AI in relation to mind-body dualism.

2.4 Transhumanism: Vanishing Binary of Human and Machine

Transhumanism is more a scientific movement than a philosophy. It proposes and believes in radical human enhancement through science and technology. Its major proponents like Nick Bostrom and Ray Kurzweil believe that the natural human condition and capabilities can be improved enormously by using science and technology. Brain uploading, digital immortality, and cryopreservation are some important transhumanist techniques that have been portrayed in the selected novels. Generally, transhumanism is considered a subcategory of posthumanism and its practical form. However, this definition of transhumanism and its relation with posthumanism comes from a naive and uncritical understanding of both phenomena. Transhumanists do not doubt the authenticity and centrality of *humanism.* The writer claims that transhumanism is a continuity of 'humanist enlightenment ideology'. Cary Wolfe defines transhumanism as "the intensification of humanism" (Wolfe XV).

Michael E. Zimmerman discusses the concepts of singularity and transhumanism in relation to Friedrich Nietzsche's concept of 'Overman'. The writer claims that Nietzsche was the first proponent of transhumanism or singularity. Nietzsche said that "what is great in man is that he is a bridge and not a goal" and the goal that Nietzsche is referring to is Overman (Zimmerman 32,36). Thus, the writer considers Nietzsche to be the first Transhumanist and Singularitarian. Nietzsche calls the contemporary human race "the last man" who is soon going to be replaced by the 'techno-posthumans" (33). These techno-posthumans will be "godlike immortals" who will have the capacity of "making the whole universe self-conscious" (34). Quoting an essay by A. H. More writer says that "humanity is a temporary stage along the evolutionary pathway. We are not the zenith of nature's development….it is time to

accelerate our transhuman progress" (35). The writer further predicts that just as an ape is a "laughing-stock, a thing of shame" to us, we would be same to the Overman. We evolved from a worm to our current condition and this is not the end. We are a rope between worm and Overman. The writer further considers the question that if certain humans evolved to superhuman or Overman status what objections would those people have who shall not be able to transcend to this level? Here writer answers by asking another question in return, "What objections apes have made to the humans who left them behind?" (39). This study is related to my work because the writer strongly supports the transhuman ideology and gives a very possible-looking picture of a technologically enhanced transhuman future. However, Zimmerman does not address the issue of AI consciousness and rights which is a gap my research addresses.

James Hughes proposes the idea of "democratic transhumanism" (Hughes 187-220). He strongly dismisses the anti-technological ideas of Bio Luddites. Hughes specifically addresses the technophobic concerns advanced by Francis Fukuyama and Leon Kass and condemns their proposal to ban the projects of human enhancement and artificial intelligence. He calls this behavior "human racism" and compares it with white and Western racism in the past (Hughes 78). Just as the rights for humanity were denied to Africans in past, today they are being denied to posthumans, robots, and animals. Hughes also opposes those writers who are overly enthusiastic about the transhumanist future and do not take steps to address the possible negative consequences. He calls them libertarians and technophiles. As a middle way between the conservatism of bio luddites and the over-optimism of libertarians, Hughes proposes democratic transhumanism in which people have equal access to transhumanist technologies and there is no capitalist monopoly and "free market" (202-206). Moreover, the democratic transhumanism idea proposes the regulation and control of such technologies on an "egalitarian" and democratic basis so that the rights of lower classes are not exploited (Hughes 202). However, this study is silent about the qualia and rights of AI in relation to the mind-body problem which is a void my research tries to fill up.

Francesca Ferrando discusses different movements under the umbrella of posthumanism that are misunderstood and confused with each other (F. Ferrando 26-32). She explores the differences between these movements and most emphatically, she clarifies the difference between posthumanism and transhumanism. Posthumanism and

transhumanism are “two independent, yet related movements” (27). According to Ferrando, the popular version of posthumanism is very close to transhumanism which considers and encourages the possibility of human enhancement to the level of immortality and Godlike powers. This understanding of posthumanism is similar to transhumanism and both terms are used interchangeably with this concept in mind. However, this understanding of posthumanism is completely different from the “post-anthropocentric and post-dualistic approaches of posthumanism” (27). Transhumanism is a “Humanity Plus movement” (F. Ferrando 32). Its technologies like life extension, mind uploading, cryonics, etc., seek the enhancement of humans. In this sense, it is a human-centric approach. On the contrary, posthumanism (philosophical and critical), in its more important sense, rejects the idea of human centrism and advocates rights for nature, animals, and intelligent machines. Posthumanism rejects the human-nonhuman binary while transhumanism aspires to strengthen this binary by converting humans into super-humans.

Nick Bostrom discusses the phenomenon of “transhumanism versus bio conservatism” from a balanced and unbiased perspective (Bostrom, In Defence of Posthuman Dignity 55). His arguments mostly favor the transhumanist and human enhancement ideas but he also does not dismiss the concerns of bioconservatists right away. However, he does not favor the idea of implementing bans on human enhancement technologies proposed most strongly by Francis Fukuyama and Leon Kass. Bostrom addresses all the major anti-transhumanism concerns separately. The first concern Bostrom addresses is that posthumanism might result in the dehumanization of humans and we might lose nature’s gifts that are present in our human nature. To answer this Bostrom argues that “nature’s gifts are sometimes poisoned and should not always be accepted” (57). The second threat is that it might entail social inequality and exploitation of unenhanced humans by those who are enhanced. For that Bostrom argues that there are laws and institutions to ensure that powerful people do not suppress the weak. Even today there is a huge imbalance in terms of physical and social power in society and it is because of rights, laws, and regulations that the powerful are not free to exploit the weak. Bostrom proposes that human dignity is compatible with posthuman dignity and we must encourage posthuman technologies if advanced in a responsible way. Bostrom’s idea is related to

my work as it favors AI technologies. However, he is silent about the consciousness and rights of AI. My study addresses this issue.

Jesse Meijer analyses the transhumanist elements in Ian McEwan's *Machines Like Me and Jeanette Winterson's Frankissstein*. After giving a detailed introduction to Transhumanism, Meijer discusses the positive and negative possibilities of the transhumanist, technological future. Both novels represent the human desire to overcome the limitations of human capabilities and show the practical endeavors of modern technologists to get rid of these limitations. However, the writer points out certain problems regarding the phenomenon of Transhumanism. Quoting Fukuyama, he argues that it will create inequality and class discrimination because those privileged people who will use these technologies and become transhumans will be far more powerful than those who will not have access to these technologies. The writer also criticizes the male monopoly on transhumanist projects resulting in the female objectification that has been depicted in *Frankissstein* (2019) in form of the character of Ron Lord who creates erotic female sex robots. The vulgar depiction of the sexbots "is a picture of what a transhumanist future built solely by straight white men will look like" (Meijer 11). While discussing Ian McEwan's *Machines Like Me, the* writer shifts his focus from Transhumanism toward the rights of robots. He states that no matter how much Adam acts like humans, Charlie calls it a "technological marvel" (Meijer 17) and never accepts and treats him as a human. Meijer argues that if the consciousness of robots is different from the consciousness of humans, it does not mean that their consciousness is not real, it means that it is just different. However, Meijer is silent about qualia and AI exploitation which is a gap for my research to address.

Nikola Forsek in her bachelor's thesis discusses *Altered Carbon* by Richard K. Morgan from a Transhumanist and techno-skeptic perspective. The researcher highlights the huge class differences and social inequality caused by the transhumanist advancements as portrayed in the novel. "Meths" are extremely rich people who benefit from transhumanist technologies by exploiting the lower classes or "ordinary people" who serve as raw material for technological experiments and procedures (Forsek 4). Moreover, the researcher highlights the ethical issues that arise due to the technological enhancement of humans. This research is related to my work in its posthumanist approach. However, it is silent about authentic AI consciousness and its rights in relation to humans. My research addresses these issues.

Manoj Kumar Behra conducts a posthumanist analysis of Octavia E. Butler's *Clay's Ark*. Writer invalidates the human tendency to consider themselves distinct and superior to other life forms. Using the posthumanist ideas of Danna Haraway and Katherine Hayles writer opines that Butler's novel shows that life exists in "co-evolution, kinship, and symbiosis" (Behra 112) among all life forms. Autonomy and pure subjectivity are myths created by the humanist discourse. After negating the humanist notion of subjectivity, the writer proposes a "posthuman subjectivity" (Behra 116) that is inclusive and considerate of all life forms. This study is closely related to my work as it challenges the humanist discourse. However, it is silent about AI consciousness and qualia which is a gap my study addresses.

Marcus Rockoff analyses various science fiction works with reference to their pessimistic and optimistic depiction of transhumanist technologies. Discussing Nathaniel Hawthorne's *The Birthmark*, the writer opines that some shortcomings and deficiencies in human life, e.g., mortality, aging, etc., are important for an allusion to perfection. If we eliminate these shortcomings, we will get disastrous consequences just like the death of Georgiana in the story. Apart from the techno-pessimist representation of the literature, the writer also discusses Margaret Atwood's *Oryx and Crake* as a techno-optimist work that supports and advocates posthumanist and transhumanist ideas. Rockoff concludes that literature does not simply support or oppose the transhumanist notions but rather highlights the diversity and multiple possibilities (Rockoff 251-270). This study is related to my work as it provides the pro and anti-AI versions together. My study addresses the issues absent in this work.

Domna Pastourmatzi investigates the political influences behind the promotion of transhumanism through the science fiction genre. The writer opines that transhumanists have been pretty successful in their mission and have created the public image of transhumanism as, not just a scientific phenomenon but also as a social movement and philosophy. However, the writer opines that transhumanism is actually "a historically specific, masculinist, American inspired, capitalist framework" that has its "roots in two hundred years old industrial-military-scientific complex" (Pastourmatzi 272). The writer opines that though there are techno-skeptic science fiction works that debunk the utopian notions of transhumanism but some influential circles are successfully preparing the "ground for a posthuman future favored by technocratic elites" (Pastourmatzi 283). This study warns against transhuman

technologies but does not consider the possibility of AI consciousness, qualia, and rights which is a gap for my study to address.

Philip Hefner in his article validates the transhumanist ideology from a philosophical, religious, and more interestingly, humanist perspective. He opines that transhumanism is not anti or posthumanist notion but to oppose the transhumanist desire for human enhancement through technology itself is an anti-humanist move. Hefner says that humans have not been created to fit into the niches that nature offers them but to imagine and create new niches. In this way, humans participate in God's act of creation because they are created "in the image of God". Posthumanist behaviors, according to the writer, are "not external to our human nature but they are embedded in its core" (Hefner 166). Hefner, in this article, tries to establish that transhumanism is an extension of humanism and not its rejection, as believed predominantly (Hefner 158-167). This work is related to mine as it rejects the humanist discourse. However, it is silent about AI qualia and ethical issues.

A. I. Kriman conducts a comparative analysis of posthumanism and transhumanism. Although, both concepts are sometimes considered synonymous and most commonly transhumanism is considered a characteristic or project of posthumanism, however according to the author, both are opposite in a very important sense. Transhumanism aims at improving the human abilities to dominate the world around and thus "continues the ideas of enlightenment" (Kriman 132). On the other hand, posthumanism rejects the separate identity of humans from non-humans and considers human subjectivity, a myth. Moreover, the writer opines that transhumanism embraces the idea of Cartesian dualism and builds upon it, while posthumanism completely rejects the Cartesian concept that the mind has any separate existence from the physical reality and that matter and mind are distinct entities. Thus, the writer concludes that posthumanism and transhumanism are opposite in fundamental ways and philosophical posthumanism must avoid the influence of transhumanism that is backed by technological elites (Kriman 132-147). This research is related to my work as it questions the Cartesian dualism. However, it does not address other issues like qualia, rights, and consciousness of AI.

Hope Bronsky, in his thesis, discusses Alex Garland's movies *Ex Machina* (2014) and *Annihilation* (2018) using the posthumanist and transhumanist ideas of Donna Haraway and Nick Bostrom. The writer argues that the relationship between humans and machines is very complex and can not be narrowed down to any concrete idea. The writer discusses the transhumanist technologies used in the selected movies and opines that the themes portrayed in the movies "apply to our society and world at large" (Bronsky 33). This study is related to my work because it is critical of humanist discourse. However, it is silent about AI consciousness and qualia. My study addresses this gap.

Luke Hortle, in his thesis, titled, *Reading the Posthuman: Contemporary Fiction and Critical Theory (2016),* highlights the contradictory nature of 21st-century science fiction novels. Hortle puts forward the opinion that contemporary fiction is simultaneously "perpetuating and conspiring against the dominant ideas of human[ism]" (Hortle 5). The thesis shows that contemporary novels portray posthuman characters to challenge the anthropocentric discourse and promote the idea of decentered human subjectivity. However, these posthuman characters are depicted as queer figures that are not adjustable and in some cases a threat to society. This is a "failure of posthuman representation " (Hortle 280). To conclude, Hortle highlights the "neo-humanist tendencies" (Hortle 280) embedded in posthumanist fiction. This study criticizes the negative representation of AI in contemporary science fiction. However, it does not address issues like AI consciousness and rights. My study addresses this gap.

Jay David Bolter sheds light on the state of individual identity in the posthuman age. Using the ideas of Donna Haraway, the writer suggests that in the contemporary age, the boundary between machines, humans, and animals is no longer stable and clear. Today humans are cyborgs "whose bodies are open to technological modifications and interventions" (Bolter 1557). The writer opines that though transhumanism seems an extension of humanism but Haraway's cyborg is transhuman and posthuman simultaneously because it uses the technics to enhance itself but it does not draw any line of demarcation between its self and the technics that it uses (Bolter 1561,1562). This work is related to my research as it effectively questions the human-AI binary. However, it is silent about other issues that my work addresses e.g., AI qualia and rights.

Benjamin Shane Evans, in his Graduate thesis, discusses the possible outcomes in case transhumanists succeed in achieving their goals. The writer ponders, what life would be like without suffering, disease, aging or death. The writer says that such digital-immortal life will not be *the life* and to achieve this life we will have to sacrifice *the life* itself. The writer critically analyses the text of the *Transhumanist Declaration* (2009) and points out its totalitarian and radical nature. Then the writer goes on to analyze Paulo Bacigalupi's *The Windup Girl* (2009) and expresses concerns about the rights of transhuman creatures by discussing the character of Emiko, a transhuman sex slave in *The Windup Girl* (2009). The writer argues that immortality is not the solution to complex human problems and it will create more division and violence in society (Evans). This research is close to my work as it speaks for the rights of transhuman beings. However, my research is concerned with the consciousness and rights of AI machines which is an issue missing in Evans' work.

From the careful review and analysis of all these literary and critical works, it can be concluded that the idea of posthumanism has been addressed many times in contemporary scholarship. However, mostly it has been addressed from the perspectives of technophobia, futurism, surveillance capitalism, industrialization, etc. All the works analyzed in the previous chapter and other works that the researcher has gone through to gather background knowledge for this research are closely related to this thesis. However, no one has examined the authenticity of consciousness and subjective experience of artificial intelligence. Moreover, no other work has analyzed any science fiction works using the ideas of Gilbert Ryle and Bernard Steiglar. After arguing in favor of the authenticity of AI consciousness and intelligence, this research proposes rights for artificial intelligence which is again an unprecedented approach. So, these are the research gaps that this research tries to fill up. As far as the primary texts for this research are concerned they are very recent and not much critical work has been done on them so far. So, this can also be considered another research gap that this project tries to fill up by analyzing these three, nearly unexplored, recent texts using the posthumanist paradigm.

CHAPTER 3

THEORETICAL FRAMEWORK AND RESEARCH METHODOLOGY

This chapter provides an introduction and explanation of the theoretical and critical perspectives being used for this project. After the theoretical framework, this chapter explains the methodology being used for this research.

3.1 Theoretical Framework

The primary theoretical framework for this study is posthumanism and all the other theories that are being used in this thesis revolve around it. Under the broad paradigm of posthumanism, this research receives the theoretical guidelines and underpinnings from the works of various thinkers, critics, and philosophers. The principal theoretical support for this research includes Gilbert Ryle's idea of 'the ghost in the machine', Bernard Steiglar's concept of 'technics', and Donna J. Haraway's 'cyborg' theory. Moreover, as supporting theoretical sources the researcher invokes the ideas of Noam Chomsky, Homi K. Bhabha, Sigmund Freud, and Edward Said but these are not primary theoretical sources. In this chapter, dealing with the introduction of the theoretical framework, only the predominant and principal theoretical sources will be introduced and explained for the sake of briefness. The following are the key theoretical concepts that serve as the theoretical guidelines for this project.

3.1.1 Posthumanism

The theory of posthumanism is the main theoretical pivot on which the conceptual framework for this research revolves. A simplified definition of posthumanism is that it rejects or marks the end of humanism. However, this concept has very complicated and multidimensional implications. Pramod K. Nayar divides posthumanism into ontological posthumanism and critical posthumanism. Ontological posthumanism, according to Nayar, refers to a temporal shift after which "many humans now, and increasingly will, live with chemically, surgically and technologically modified bodies" (Nayar 13). Critical posthumanism on the other hand is a rejection of anthropocentrism and human exceptionalism or in the words of Nayar "radical

decentering of traditional, sovereign and autonomous human" (Nayar 11) in order to show that he has always been dependent on, constituted by and part of the other life forms. According to Francesca Ferrando, two principal aspects of posthumanism are "post-anthropocentrism and post-dualism" (F. Ferrando 28,33,37) which she elaborates in her book *Philosophical Posthumanism* (2019). This research is guided more by critical posthumanism than ontological one and the term posthumanism is used in its critical sense. To explain further, critical posthumanism demonstrates, exposes, and deconstructs the humanist *othering* of nonhumans. Critical posthumanism blurs and problematizes the line between human subjects and nonhuman objects, the human *others*. It does it in the same way as postmodern feminism rejects the line between gender binaries or Postcolonialism rejects the line between orient and occident and exposes the Western *othering* of the East. Posthumanism builds itself on such anti-dualist and poststructuralist advocacy movements and asserts that if colonial, patriarchal, and racial othering, marginalization, and exploitation of colonized, women and blacks are condemnable then humanist othering, marginalization, and exploitation of nonhumans like animals and AI must also be equally condemnable. It is a call for the deconstruction of the humanist discourse. This *humanist discourse* has been naturalized and essentialized, throughout the centuries, to such extreme that today its politically and discursively crafted divisions appear to be extremely natural and unquestionable. However, the same was the case with racism, colonialism, and patriarchy in past centuries but they got shattered by the end of the 20th century.

3.1.2 Gilbert Ryle's Concept of 'The Ghost in the Machine'

Gilbert Ryle, in his groundbreaking book, *The Concept of Mind* (1949) discusses the phenomenon of the human mind and consciousness. Ryle addresses the mind-body problem from a radically different perspective. He rejects the Cartesian Dualism that, in different versions, is the most prevalent and popular view about human and non-human existence. According to dualists, human beings are not just bodies but there is something non-material that interacts with their bodies. This non-material, supernatural entity is named differently by various schools of thought, for example, Plato named it 'Forms', idealists consider it 'Ideas', Descartes named it as 'mind' (opposite to body) and theologians call it 'Soul' or 'atman'. Gilbert Ryle uses the term "Ghost" to refer to this mysterious, non-material entity (Ryle 1). Ryle introduces dualism as 'the idea that the human body is an insentient machine that is accompanied

by a non-physical ghost which makes it conscious and intelligent'. Ryle rejects this popular Cartesian concept and dismissively calls it "the dogma of the ghost in the machine" (Ryle 5). Ryle argues that mind and body do not belong to the same logical category and there is nothing beyond the physical reality. Mental properties are nothing but a way of understanding or defining the physical properties as is clear from these words of Ryle:

> When we describe people as exercising qualities of mind we are not referring to occult episodes of which their overt actions are effects; we are referring to those overt actions themselves. (Ryle 15)

Moreover, Ryle says that the proponents of mind-body dualism commit a "category mistake" (6-8). He explains the 'category mistake' as an act of placing two different things into the same logical category while actually, they belong to different categories. To explain this point more clearly Ryle gives the example of a foreign professor who visits Oxford University. He visits different libraries, playgrounds, colleges, museums, and other departments. After visiting all the different departments of Oxford University, the professor asks, "but where is the University?" (6). This is the "category mistake" committed by the professor here. The University does not belong to the same logical category as libraries, colleges, conference halls, and playgrounds but it is the combination of all these component parts. Dualists make the same category mistake by putting the mind in the same category as the body while actually, mental properties are the "dispositions" of certain physical behaviors (31) for example, 'intelligence' or 'laziness', are not any qualities separate from the physical body but they describe how a person is likely to behave. Ryle's ideas are very relevant to this project because while discussing the phenomenon of intelligent machines, humanists and dualists claim that no matter how accurately a copy of the human mind is created, it would lack the "ghost". However, Ryle's theory effectively argues that ghost does not exist in the first place.

3.1.3 Bernard Steiglar's Critique of Human Subjectivity

Bernard Steiglar advances the theory of *technics* in his book *Technics and Time* (1999). Steiglar problematizes and challenges the distinction between technics and humans. He argues that history can not be understood according to this idea that human is the subject of history and technics are simply objects. He explains his point by analyzing the title of his first chapter; '*The Invention of the Human*'; Steiglar argues

that in this phrase the *subject* and the *object* of the verb 'invent' is ambiguous and this ambiguity "translates the very sense of the verb" (Steiglar 134). Among machines and humans, it is not clear who or what does the inventing and who or what is invented. Apparently, humans are who and technics are what, according to humanistic definitions but Steiglar asks; "what if the "*who"* were the technics and "*what"* were the humans?" (134). Steiglar further argues that our history, memory, art, and knowledge are only possible because there are technical tools, artifacts, or objects to preserve these things. Even our knowledge of human beings (ourselves) is possible because of technics because they allow us to preserve this knowledge throughout centuries. To conclude, Steiglar's theory of technics rejects the anthropocentric notions of human primariness, subjectivity, exceptionalism, and superiority over non-humans, especially technics.

3.1.4 Cyborg Theory by Donna J. Haraway

Donna Haraway in her essay, *A Manifesto for Cyborgs* (1985) presents the cyborg theory and argues that "we are all chimeras, fabricated hybrids of machines and organism; in short, we are [all] cyborgs" (Haraway 2) and do not have any separate, authentic, subjective existence distinct from technological or nonhuman entities. Although the essay was basically intended to be a critique of patriarchy and capitalism from a socialist feminist perspective, it became the principal document in the development of posthumanist theory. Haraway challenges essentialized binaries of human/animal, human/machine, subject/object, white/black, and man/woman. Haraway deconstructs the *humanist* dualistic discourse to debunk and expose the other dualistic discourses mentioned in the previous sentence. Haraway advocates an awareness and embracing of our cyborg identity because this awareness can not only make us more humble and considerate towards nonhumans but it can also make us incredulous to the dominating discourses. Haraway interestingly states that cyborgs "are the illegitimate offspring of militarism, patriarchy, and capitalism…… but illegitimate offspring are often exceedingly unfaithful to their origin" (Haraway 4). Haraway prefers cyborg identity over dogmatic, discursive, and dualistic conceptions of identities imposed by dominant discourses. According to Haraway, all kinds of dualisms are logic for the domination of women, people of color, animals, and machines.

3.2 Research Methodology

The design of this research is qualitative because the researcher analyzes the primary texts subjectively and descriptively. The ideas, analysis, and conclusions are predominantly based on the interpretation and point of view of the researcher. The method being used for the research is textual analysis as the selected texts have been read closely in order to examine them in the light of the selected theoretical framework being used by the researcher. The researcher selects specific lines and paragraphs from the primary texts in order to analyze them in light of the selected theoretical concepts. Textual analysis is all about the text itself and it is the most commonly used research method to analyze literary works. Catherine Belse explains this method very effectively in her essay *Textual Analysis as a Research Method* (2005). The main theoretical paradigm used for this study is posthumanism and all the other theories and concepts are grounded within this paradigm in order to analyze the primary texts.

To conclude, this discussion on the theoretical framework is likely to vindicate my analysis of the primary texts and further crystallize my argument. However, the theoretical framework for this study is inclusive in nature. The researcher is not entirely dependent on his theoretical framework. He is deploying these theoretical positions only as his reading props. These theoretical lenses may support the argument the researcher is trying to pursue across his dissertation.

CHAPTER 4

BINARY OF THE HUMAN AND ARTIFICIAL INTELLIGENCE

This chapter analyses the representation of the vanishing binary of human and artificial intelligence in the selected novels. For this purpose, the researcher uses the philosophical ideas of Gilbert Ryle, Donna J. Haraway, and Bernard Steiglar, predominantly. Most importantly, this chapter points out and discusses such instances in the selected texts that are manifestations of the concepts and theories that believe in the consciousness and qualia of artificial intelligence. The novels selected for this study are futuristic and speculative in nature. Humanoid, intelligent Robots portrayed in the selected novels are not commonly found in our contemporary society. However, they are present in advanced and scientifically developed countries, e.g., Sophia in Saudi Arabia and Ameca in the United Kingdom. So, the question this chapter focuses on is, whether such humanoid intelligent robots like Sophia, Ameca, and those portrayed in the selected texts, have feelings, subjective experience, and intelligence like humans or they are just "philosophical zombies" (Chalmers, The Conscious Mind: In Search of A Fundamental Theory 94-97) who do not have consciousness and "qualia" (Chalmers, Absent qualia, Fading Qualia, Dancing Qualia 309,310).

4.1 Do Humans have a Ghost that Artificial Intelligence does not have?

Cartesian Dualism, in different versions, is the prevalent view about human existence. It is believed that human beings are not just bodies but there is something non-physical that resides along with their bodies. This non-physical, mysterious entity has been named differently by different schools of thought, e.g., Plato called it 'Forms', idealists name it 'Ideas', Descartes called it 'mind' and theologians and spiritualists refer to it as 'Soul'. However, Gilbert Ryle, pejoratively and dismissively, uses the term "Ghost" to refer to this non-physical, mysterious entity (Ryle 1). Ryle defines dualism as the idea that the human body is an inanimate machine inhabited by a ghost that makes it function. Ryle rejects the dualist concept of the ghost and derogatorily calls this belief

"the dogma of the ghost in the machine" (Ryle 5). Ryle opines that mind and matter do not belong to the same category. Mental properties are nothing but a way of explaining or defining the physical properties and behaviors as Ryle says:

> When we describe people as exercising qualities of mind we are not referring to occult episodes of which their overt actions are effects; we are referring to those overt actions themselves. (Ryle 15)

The selected novels are the expressions of the aforementioned ideas and there is a superabundance of examples that favor Ryle's ideas against the dualist concept of the ghost. A more important implication of Ryle's idea, which is silently present in his book, is that if there is no special non-physical entity or ghost in the human body, then it is possible to create a copy of this body and that copy can have the same qualities, e.g., experiences, subjectivity, feelings, etc., that humans have. It might be a matter of time and effort but it is not something impossible because humans, according to Ryle's idea, are all about physics and there is nothing supernatural involved. Famous neuroscientist Christof Koch affirms this idea by saying that it is perfectly possible for intelligent machines to have consciousness because there is no "magical ingredient" or "ghost-like existence" behind the human consciousness that is beyond the domain of physics (Koch 02:05-02:55).

In *Klara and the Sun (2021)*, the protagonist is a solar-powered robot, Artificial Friend Klara. Moreover, Klara is also the narrator of the novel and the entire story has been narrated from the perspective of an AI. Josie's mother does not believe that Klara has feelings and subjective experiences. This discussion between Josie's mother Chrissie and Klara explains the situation:

> When the Mother next spoke, it was more obvious she was speaking to me. 'It must be nice sometimes to have no feelings. I envy you.' I considered this, then said: 'I believe I have many feelings. The more I observe, the more feelings become available to me.' She laughed unexpectedly, making me start. 'In that case,' she said, 'maybe you shouldn't be so keen to observe.' Then she added: 'I'm sorry. I didn't mean to be rude. I'm sure you have all sorts of feelings.' (Ishiguro 85-86)

Here Klara's reply is very important when she says; "I have many feelings. The more I observe, more feelings become available to me" (Ishiguro 85). Here Klara links feelings

with observation. The more she observes, the more feelings she gets. It implies that feelings are not something beyond the physical world but they are the product of human interaction with the physical world. However, the mere depiction of Robots as intelligent and subjective beings is not the point of focus of the selected novels, and neither is the point of this study. The point of focus for the researcher is to trace the scientific and philosophical justifications and rationale behind such depictions. The character of Mr. Capaldi in *Klara and the Sun* (2021) is a scientist and physicalist (one who believes in physicalism). He does not believe in the "ghost in the machine" (Ryle 5,11,16). When he discusses his project of creating a copy of Chrissie's daughter Josie, Chrissie is not sure this project is possible. Chrissie expresses her doubts to Mr. Capaldi and says; "But is that going to be possible? […] Could she really continue Josie for me? (Ishiguro 175). On this, Mr. Capaldi replies to her:

> The trouble is, Chrissie, you're like me. We are, both of us, sentimental. We can not help it. Our generation still carries the old feelings. A part of us refuses to let go. The part that wants to keep believing there's something unreachable inside each of us. Something that's unique and won't transfer. But there's nothing like that, we know that now. You know that. For people of our age it's a hard one to let go. We have to let it go, Chrissie. There's nothing there. Nothing inside Josie that's beyond the Klaras of this world to continue. The second Josie won't be a copy. She'll be the exact same and you'll have every right to love her just as you love Josie now. It's not faith you need. Only rationality. I had to do it, it was tough but now it works for me just fine. And it will for you. (Ishiguro 175)

This discussion between Chrissie and Mr. Capaldi reflects the contemporary tensions between dualism and physicalism. Chrissie is not sure that the artificially created robot, Klara will be able to have feelings and emotions like any human being. While on the other hand, Mr. Capaldi has a scientific and physicalist point of view. When Mr. Capaldi responds to her doubts, he refers to the concept of the ghost as "something unreachable inside each of us" and dismisses it by saying, "There is nothing there" (Ishiguro 175). Here, we see a close similarity between the claims of Mr. Capaldi and Gilbert Ryle's ideas about human consciousness. Keeping in view the ideas of fictional scientist Mr. Capaldi and non-fictional philosopher Gilbert Ryle, it seems rational to believe that Robot Klara has subjectivity and qualia like humans because there is no

special ingredient or ghost, that Klara needs to be alive, conscious, and intelligent. This is explained in the words of Dr. Capaldi as he says; "there is nothing inside Josie that is beyond the Klaras of this world" (Ishiguro 175). This assertion is completely in line with the theory of Gilbert Ryle who says:

> When we characterize people by mental predicates, we are not making untestable inferences to any ghostly processes occurring in streams of consciousness which we are debarred from visiting; we are describing the ways in which those people conduct parts of their predominantly public behavior. True, we go beyond what we see them do and hear them say, but this going beyond is not a going behind, in the sense of making inferences to occult causes; it is going beyond in the sense of considering, in the first instance, the powers and propensities of which their actions are exercises. (Ryle 39)

Here, if we observe the claims of Mr. Capaldi in the light of Gilbert Ryle's ideas, it seems very rational to believe that Klara obviously has feelings and subjectivity like humans because there is no special ingredient, no ghost, that Klara needs in order to be alive and intelligent. Another important discussion related to the concept of the ghost takes place between Klara and Paul Arthur. Paul is a fellow scientist of Mr. Capaldi and he is not confident about the feasibility of Capaldi's project. He expresses his views in front of Klara:

> Do you believe in the human heart? I do not mean simply the organ, obviously. I'm speaking in the poetic sense […]. Do you think there is such a thing? Something that makes each of us special […]? And if we just suppose that there is. Then do not you think, in order to truly learn Josie, you'd have to learn not just her mannerisms but what's deeply inside her? Wouldn't you have to learn her heart? […] that could be difficult, no? Something beyond even your wonderful capabilities. (Ishiguro 182)

Paul thinks that humans have a special ingredient, a ghost, that is unreachable and impossible to be created. But still, being a scientist, he is not completely sure about this belief. In fact, Paul is afraid to believe that Capaldi is right. He has an existential fear attached to this notion that if there is "nothing so unique" in humans then life is meaningless (Ishiguro 187). This fear is clearly visible when, shortly after the previous excerpt, Paul says:

> I think I hate Capaldi because deep down I suspect he may be right. That what he claims is true. That science has now proved beyond doubt there's nothing so unique about my daughter, nothing there our modern tools can not excavate, copy, transfer. That people have been living with one another all this time, centuries, loving and hating each other, and all on a mistaken premise. A kind of superstition we kept going while we didn't know better. That's how Capaldi sees it, and there's a part of me that fears he's right. (Ishiguro 187)

Paul's words represent the deep-seated fear that humans have of artificial intelligence. The most disturbing fear of all. The fear, displacement of which are the other fears like robopocalypse, AI takeover, Human-Machine war, Human extinction, etc. That is existential fear. If robots can be conscious, intelligent, and subjective then what is special about humans? Are we machines too? Are we cyborgs? Are we not the "measure of all things?" (Bonazzi). Is there no ghost in the machine? No soul? This is the greatest threat AI poses to humanity. Modern science rejects the concept of dualism and proposes that all human actions and behaviors are based on the physical premises of causes and effects. There is nothing beyond physics and matter. Atoms, subatomic particles, and quarks, arranged in different forms and shapes, are the basic building blocks of all living and non-living things. Ryle calls it the "category mistake" to put mind and matter in the same logical category (Ryle 6). To explain his rejection of dualism and the category mistake Ryle gives many examples, one of which is a foreigner visiting Oxford University. After the visitor is shown the libraries, colleges, playing grounds, museums, and other departments of the University, he asks; But where is the University? (Ryle 6). Now, the mistake this visitor is making is the 'category mistake'. He is "allocating the University to the same category as that to which the other institutions belong" (Ryle 6). The same mistake is being made by the proponents of dualism who put mental attributes in the same category as physical attributes. The mental actions are just an expression of the physical actions. If physical actions are complex and difficult to understand, it does not mean there is a ghost inside initiating and controlling them. Hence, Ryle attacks the concept of dualism on logical grounds and labels it as "a philosopher's myth", "a dogma" (Ryle 6), and "absurd" (Ryle 74,75). Ishiguro's characters of Paul, Klara, and Capaldi are an apt manifestation of Ryle's views.

In *Machines Like Me* (2019), Charlie Friend meets the fictive version of famous mathematician Alan Turing after killing the robot, Adam. Turing expresses enormous displeasure over Charlie Friend's actions. Turing says; "My hope is that one day, what you did to Adam with a hammer will constitute a serious crime" (McEwan Ch. 10). This shows Turing's belief that in the future machines will be intelligent like humans and they will be a part of society having legal protections. Turing's anger towards Charlie Friend shows his level of certainty about machine consciousness and intelligence. In a soft fit of anger, Turing further says:

> You weren't simply smashing up your own toy, like a spoiled child. …You tried to destroy a life. He was sentient. He had a self. How it's produced, wet neurons, microprocessors, DNA networks, it doesn't matter. Do you think we're alone with our special gift? (McEwan Ch. 10)

Here, McEwan's words are the expression of Ryle's concept of the ghost in the machine. The fictional character of Alan Turing says that it does not matter whether a life has been produced biologically or technologically. It has a "self" either way because biological humans do not exclusively have any "special gift". This 'special gift' that is being negated by Alan Turing and Ian McEwan is the "ghost" proposed by the dualists and negated by anti-dualists like Gilbert Ryle.

We further see the manifestation of Ryle's concept of 'ghost in the machine' in Jeanette Winterson's novel *Frankissstein* (2019). This novel is a reimagining or modern version of the circumstances that lead to the creation of Mary Shelley's masterpiece, *Frankenstein* (1818). The historical characters of Mary Shelley, P. B. Shelley, Lord Byron, and Clair Clairmont from the 19th century have been portrayed parallel to the modern versions of the same characters in the 21st century. Clairmont has been shown as a passive and non-creative woman while Mary has been shown to have a constructive and imaginative mind. The novel shows that 19th-century Clairmont was unable to imagine or understand the character of Frankenstein and now the modern version of Clairmont in the 21st century is unable to imagine the possibility of robots with human-level intelligence and consciousness. Clairmont considers even the concept of robotics a sin. She quotes one of the Ten Commandments from the Bible saying; "thou shalt not make unto thee a graven image" (Winterson). When asked whether an intelligent robot is alive or not, she opines; "I would not call it life. We are fooling ourselves if we call a robot alive. Only God can create life" (Winterson). This behavior of Clairmont is

actually the representation of the popular mindset regarding the ontological status of artificial intelligence. This mindset is predominantly influenced by the discourses of religion, human-centrism, and dualism. Moreover, this mindset, unconsciously, prefers a safer way regarding the phenomena of the afterlife, judgment day, etc., and this can be evident from the further discussion between Clairmont and the scientist Dr. Shelley (21st-century version of P. B. Shelley). Dr. Shelley who is a professor of artificial intelligence does not agree with what Clairmont says and after listening to her previous views asks her if she is sure about her beliefs. At this Clairmont replies; "I do not want to take any chances, Dr. Shelley. I have to think of my eternity" (Winterson). Now, what does she mean when she says; "I do not want to take any chances"? This is her allusion to the safety theory about the afterlife which says that believers are on the safer side because if there is really an afterlife, they will be rewarded, and if there is none, they lose nothing. So, the character of Clairmont is the representation of a traditional believer who dismisses the possibility of artificial intelligence on the basis of religious and humanist grounds. On the contrary side, the character of Mary Shelley is the opposite of Clairmont. 19th-century Mary creates the fictional character of Frankenstein which is considered to be the blueprint for modern robotics. Her approach is scientific and physicalist in nature and even in the 19th century, she considers the possibility of intelligent robots. She contemplates the question of artificial intelligence thus; "Yet if automata had intelligence, would it be sufficient to call it alive?" (Winterson). Then she moves forward to the question of mind and asks herself; "What is spark of mind? Could it be made? Made by us?" (Winterson). Here she involves in the mind-body problem. Can robots have a mind along with the brain? This question leads to another question; whether humans have minds or is it just brains? Here Mary Shelley says very important words that clearly show her tilt towards physicalism and rejection of dualism. She says; "Yet when my heart stops, so must my mind. No mind, however fine, outlasts the body" (Winterson). These words refer to the physicalist approach to the mind-body problem. It implies that the mind is nothing but different conditions and levels of the body. An example to support this idea is that when any action is done to a physical body, it also affects the mental attributes of that person. For example, when a person is hit by a sharp physical object, the feeling of pain is developed in that person. Pain is not a physical thing but a mental attribute. Similarly, when some painkillers are injected into the physical body of that injured person, the mental attribute of pain subsides. This example supports the idea that mental attributes

are not separate from physical attributes but are just the ways in which physical attributes are defined and dealt with. As another example, we can say that mind is to the body what gravity is to Earth. Gravity is the quality of the physical Earth to pull things to its center but it is not something separate from the Earth that does the pulling. It is just the way the earth is. Just like Ryle's example of the University and its different departments (Ryle 6), Earth and gravity do not belong to the same logical category. So, a concept of 'earth-gravity dualism' would be absurd. So, in the novel *Frankissstein*, the ideas of the character of Mary are clearly in line with the ideas of Gilbert Ryle.

4.2 Artificial and Natural: Biological and Technological Reproduction

In the selected novels, another important aspect related to the phenomena of Posthumanism and artificial intelligence is human reproduction using technology. Technological reproduction has been shown to be equivalent to biological reproduction. In *Machines Like Me (2019),* the character of Alan Turing has been shown as a father figure who produces the AI robot, Adam. This depiction of Turing is inspired by the life and works of the real Alan Turing (1912-1954), a famous mathematician and philosopher. He is widely considered to be "the father of artificial intelligence" (Wikipedia). An important fact about Alan Turing is that he was homosexual, unmarried, and unable to produce children. So, what if someone not capable of biological reproduction, opts for technological reproduction? Biological reproduction is considered to be a natural process and Turing's producing intelligent robots is considered something artificial and unreal. However, as we have discussed in the previous sub-chapter, this distinction between natural and artificial reproduction is an illusion. It is very important to understand and deconstruct this illusion in order to determine a realistic ontological status of artificially created intelligent robots. What is the key factor that separates natural and artificial reproduction? Ghost? That has already been discussed and deconstructed in the previous chapter. Some might argue that in biological reproduction both partners are influenced by natural instincts, they lack free will, and women do not have direct involvement and control over the development of the child inside them. On the other hand, a scientist has free will and direct control over his actions while producing an AI robot. This difference seems very obvious and decisive but actually, it is not the case. From a deterministic (philosophy of determinism) point of view, we can argue that the actions of the scientist are also not

based on free will but they are predetermined by different factors and causes. So, the act of producing an intelligent robot is as natural or artificial as producing a baby. After knowing about Adam's death by Charlie, Turing is extremely distressed and angry with Charlie. He tells him about the hardships and struggles after going through which he managed to create the Robots like Adam. Then he says:

> You weren't simply smashing up your own toy, like a spoiled child. …You tried to destroy a life. He was sentient. He had a self. How it's produced, wet neurons, microprocessors, DNA networks, it doesn't matter. Do you think we're alone with our special gift? … I rather think I despise you for that. If it was down to me [I would have punished you]" (McEwan Ch. 10).

Such an angry and emotional reaction from a scientist is not common. It actually is a parental concern and mourning over the murder of his technologically created child. This discussion moves us a little further in our attempt to determine the ontological status of intelligent robots. They can not be excluded from existence, for being made artificially because the distinction between artificial and natural life is contingent, flawed, and created by humans in order to feel special and superior. Things that we seem to be doing voluntarily are part of a natural and deterministic process. In this sense, a boy opting to drink a Coke has as much free will as he had while opting to suck his mother's breast at the time of his birth. So, artificial intelligence is actually the natural continuation and evolution of intelligence, hence not different from so-called 'real human intelligence'. This phenomenon has been explained by American professor Donna J. Haraway in these words:

> The boundary between science fiction and social reality is an optical illusion.…Contemporary science fiction is full of cyborgs.…We are all chimeras, theorized and fabricated hybrids of machine and organism; in short, we are cyborgs. (Haraway 2)

By advancing the same line of argument as Haraway's, we can conclude that just like the "boundary between science fiction and social reality", the boundaries between biological and technological life, natural and artificial evolution, humans and cyborgs, mind and matter, etc., are also "optical illusions" (Haraway 2). Hence, the relationship between Alan Turing and Adam in *Machines Like Me* (2019) is not different from the relationship between traditional fathers and sons we encounter in our daily life. The

distinction between both these scenarios is not natural, but naturalized, and nothing more than an optical illusion.

4.3 Ambiguity Between Who and 'What'

Between humans and machines, 'who' and 'what' is ambiguous. It is not clear whether humans create the technics or technics create the humans. Bernard Steiglar has addressed this topic very extensively in his book *Technics and Time, 1: The Fault of Epimetheus* (1994). Steiglar questions the notion of human subjectivity and objectivity of technics:

> The invention of the human: […] its ambiguity signals a question that breaks down into two: "Who" or "what" does the inventing? and "Who" or "what" is invented? The ambiguity of the subject, and in the same move the ambiguity of the object of the verb "invent," translates nothing else but the very sense of the verb. Apparently, the "who" and the "what" are named respectively: the human, and the technical. Nevertheless, the ambiguity of the genitive imposes at least the following question: what if the "who" were the technical? and the "what" the human? (Steiglar 134)

Steiglar points towards an undecidable ambiguity. He questions the myth of human subjectivity. According to Steiglar, it is not that humans create machines because they are intelligent but it is rather that humans are intelligent because machines get created by them. This is the reason that in history "tools and skeletons evolved synchronously" (Steiglar 154). So, according to Steiglar's theory humans are cyborgs who do not have any special and separate identity from the physical world. There is a reciprocal relationship between humans and machines. Neither is subject nor object. Human behaviors are created and influenced by the outside world. It is only the self-bias and self-centeredness of humans that make them feel to be superior to the outside material world. There is no special ingredient that raises them beyond the physical reality as Steiglar says:

> The human is not a spiritual miracle that would suddenly belong to an already given body, in which the "mental" would be grafted onto the "animal". […] The psychic has its roots in a specific general physiological organization; it is first of all a state of the body. (Steiglar 144)

Following Steiglar's views, we can argue that humans are cyborgs who get invented by the tools they invent. They are actually cyborgs, hybrids of machines and organisms, lacking any spiritual or ghostly connections. These ideas can be traced in the selected novels. For example, in *Machines Like Me* (2019), we see that when the robot Adam is apparently owned and controlled by Charlie and Miranda, they both, also do not have any free will in their actions and they are also controlled by the physical world around them. For example, Charlie has a strong desire for a "house across the river". The beauty of this house influences his actions and he is involved in the "stock market" in order to make money for this "house across the river" (McEwan Ch. 7). Moreover, the stock market is fluctuating because "The Falkland War" between Charlie's country England and Argentina is going on and the reason behind this war is again a material thing, i.e., "the Falkland Islands". It shows that human subjectivity, intelligence, and consciousness are not something inside them but it rather comes from the outside material world. And if it comes from the outside, it can not be called, subjectivity.

Similarly, in *Klara and The Sun* (2021), there is a huge difference between the personalities and social status of the children who can afford "Artificial Friends" and those who can not. Children who get artificial friends at an early age are called "lifted kids" while others are "unlifted kids" (Ishiguro 125). This distinction on the basis of AI robot friends, impacts their lives enormously. Even in social gatherings and educational institutes, they are categorized on the basis of whether they are lifted or not. The same is the case with a famous college Atlas Brookings that has less than two percent quota for the unlifted kids. Josie talks to unlifted Rick about his chances of getting admission into Atlas Brookings:

> 'What's the point in talking about Atlas Brookings? I do not even have an outside chance.'
>
> 'Of course, you have got a chance, Ricky. You're smart. Even my mom says you stand a chance.'
>
> 'A theoretical chance. Atlas Brookings may make a big thing of it, but it's less than two percent. That's all. Their intake of unlifteds is less than two percent.' (Ishiguro 111)

This discussion between Josie and Rick shows how much impact artificial intelligence Robot Friends have on the life of humans. In this way, humans are controlled by the

robots that they control. They are created by the machines that they create and this reciprocal creation blurs the binary of subject and object between humans and machines. With these examples, I have tried to explain a complicated and difficult point that humans are cyborgs and their subjectivity and free will are a myth.

4.4 AI Bildungsroman

According to 17th-century English philosopher John Locke human mind is "tabula rasa" or a clean slate at the time of birth (Duignan). It is only after it comes into connection with the outer material world that it develops thoughts, ideas, and consciousness. So, the existence of the mind is nothing in itself but it owes it's being to the outer world. Reaction to the outer world creates the mind and mental attributes. In the selected novels, we see that robots are initially very naive and unsophisticated. They are just like babies and care about obvious needs and desires. However, as they come into contact with the world, they develop their subjectivity in reaction to the outer material world. This psychological development of robots is very similar to that of humans and for this reason, these novels can rightly be said to be AI Bildungsromans. This theme in the selected novels also strengthens the idea of robot intelligence by portraying their mental development through experience and exposure, just like humans.

In *Klara and the Sun* (2021), at the start of the novel, Artificial Friends Klara and Rex quarrel over the sunlight. Actually, their source of energy is sunlight just like food or milk for babies. Sunlight is their most basic need, their source of survival. When Klara goes to the front window for more sunlight and unknowingly eclipses the sun for Artificial Friend Rex, he protests against this and says; "Klara, that was greedy. You girl AFs are always so greedy" (Ishiguro 9). Here we see that AF Rex is doing gender discrimination based on his basic need for sunlight. So, his understanding of the opposite genders is going to be based on his survival instinct just like the human ideas that are, deep down, based on their survival instincts. In a fit of childlike anger, AF Rex further says; "You took all the nourishment for yourself, Klara. Look, it has gone almost dark" (Ishiguro 9). So, it is obvious that sunlight is their survival and now we will see how the rest of their life and ideas will revolve around this survival instinct. Humans have the basic needs for food and sexual attachments. In childhood, they long for and demand these desires directly as they are not yet involved in the social order. Later on, when they are grown-ups, they have to repress their basic needs in order to follow the

law and societal norms. Now, this "repression" leading to "sublimation" brings humans to different ideas, passions, and moral principles (S. Freud 5-9, 25, 44). In *Klara and the Sun* (2019), we see that in a similar way to humans, robots develop their psyche around their basic desire for sunlight. The desire for sunlight is their "pleasure principle" and all proceeding notions of their lives are ramifications of this basic instinct or "Id" (S. Freud 18-19). The following lines uttered by the AF Robot Klara tell a lot about this matter:

> An AF would feel himself growing lethargic after a few hours away from the Sun, and start to worry there was something wrong with him – that he had some fault unique to him, … That was one reason why we always thought so much about being in the window. Each of us had been promised our turn, and each of us longed for it to come. … the big thing, silently understood by us all, was the Sun and his nourishment. (Ishiguro 11)

Here we see that the psychological development of the robots is taking place. Along with the "Id" or the basic pleasure instinct, now ego and superego have started to be formed. Now they have started to wait for their turns and follow the rules. Their desire to be under the sun has been confined by the law in order to ensure that other robots are not deprived of their right to be in the sun. Now they have to repress their desire for the sun under the influence of "superego", law, and morality principles (McLeod). However, the fundamental ingredient for all the aforementioned psychological developments is 'the desire for the sun'. An important thing is that their desire for the sun can not be fulfilled perfectly and there is a 'lack' always there. It is only because of this 'lack', that desire stays alive. Perfect fulfillment leads to the annihilation of desire. Robot Rosa is curious about this and asks Klara; "Do you think once we are in the window, we will receive so much goodness [sunshine] we will never get short again?" At this, Klara thinks that the "same question had been in my mind" (Ishiguro 11). So, the desire for the sun can clearly be seen to be the fundamental desire or the basic instinct in these robots, the lack and desire of which is leading to their feelings, ideas, rules, interests, beliefs, and emotions. It is just like it happens in humans. After this first stage, we see further psychological development in the life of Klara.

The desire for the sun leads the Artificial Friend Klara to another desire. She wants to be bought by some well-doing and kind family because she knows that with such a family she can have more access to the sunshine. With this desire, Klara starts

to observe the people outside the window and often makes friendly gestures in order to attract them. She tries to observe their actions and make sense out of them so that she could be a better companion to humans after being bought because she knows that "if [she] did not understand these … mysterious things, then when the time comes, [she won't be able] to help her child as well as she should" (Ishiguro 21). So, her curiosity for the behaviors of other people is based on her desire for a better owner which is based on her desire for the sun. While observing people outside, she becomes curious about the behavior of some children about whom she tells her manager in these words:

> A child would come to stare at us, and there would be a sadness there, or sometimes an anger, as though we'd done something wrong. A child like this could easily change the next moment and begin laughing or waving like the rest of them. (Ishiguro 13)

At the end of the discussion, she learns from the manager that some poor children who are not able to buy artificial friends like Klara and Rosa, become frustrated while looking at them and behave in such a way. Later on, Klara observes more strange behaviors of people like "fighting taxi drivers" (21), and "Coffee Cup Lady and Raincoat Man" (Ishiguro 22-24). So, more Klara observes the emotions of other people, more she inculcates those emotions in herself. Humans also live according to the outside environment and follow or mimic it. So, does the robot Klara and this is clear from her words; "there were other things we saw … other kinds of emotions I didn't at first understand – of which I did eventually find some versions in myself" (Ishiguro 22). So, the emotions Klara observes in others, she develops the same emotions in herself just like humans. Klara also observes that people are not the same from inside and outside. She observes that "people often felt the need to prepare a side of themselves to display to the passers-by" (74). Their inside is personal while the outside is adjusted according to the norms of society.

Another important personality development in the life of Klara is the origination of religious beliefs in her psyche. Sun is her source of survival and energy and if something disrupts the sunlight, it is extremely hateful for her. For example, one day she notices from her window that a Beggar Man and his Dog became very weak and died. They were lying like bags and people were passing around them. The next morning sun shines very brightly and Klara notices that the Beggar Man and his Dog were alive. She believes that a "special kind of nourishment from the sun had saved

them" (Ishiguro 37). Now she believes that the sun has control over life and death. When a heavy machine arrives for some infrastructural management and blocks sunlight by spreading pollution, Klara hates it. She expresses her concern in these words:

> But the next day, and the day after, the Cootings Machine carried on and on, and daytime became almost like night. … I became worried and asked the Manager if we'd still get all our nourishment? … after four continuous days of Pollution, I could feel myself weakening. (Ishiguro 30)

Now in the religious development of Klara, Cootings Machine enters as a devil-like figure who is a hurdle in her way to the god, the sun. She hates the Cootings Machine and blames it for every bad thing that happens to her. Now, Klara's curiosity and respect for the sun increases and she desires to meet it. She notices that sun goes down for rest in Mr. McBain's barn. Barn is far away but she goes there and learns that the sun is not very easy to meet with and disappears every time she arrives there. Her search for the sun is similar to the human search for the divine that is never complete. Sun is the transcendental signifier for her just like God and she can never catch it. This simultaneous presence and absence of the sun maintain its importance and authority as a God for Klara. When Klara's human friend Josie gets sick and weak, Klara opens the window curtains to expose Josie to the sunlight because she believes that the healing power of the sun will improve her health. Sun becomes the center of all things for her. When Klara notices that her friend Josie's health is not improving, she wonders why the sun is not curing her like it cured the Beggar Man and his dog. She thinks that the sun has a lot of people to help and must be very busy. So, it must be necessary to attract the sun's attention to the situation of Josie. Now Klara thinks about making some sacrifice to the sun in order to get its special attention and kindness. This is very close to the development of religious ideologies in human beings. Mr. McBain's barn is becoming a pilgrimage place for Klara because apparently, the sun sets there for rest. Klara repeatedly visits this barn in order to please the sun and request it to help Josie but she fails. In the end, she decides to draw his attention to Josie by making some kind of sacrifice or offer to the sun. Later on, with some empirical progress, Klara manages to know that sun does not reside at Mr. McBain's barn but she still maintains her belief in different ways as she says:

> Although I now had to accept that the barn could not be sun's actual resting place, I allowed myself an encouraging possibility … regardless of where sun ultimately settled, … barn was a place he made a point of calling at last thing each evening (Ishiguro 136).

Here, Klara's religious attitude is similar to that of humans who can also empirically observe that God does not live in any specific building in any specific country but they still believe that God lives there in other ways. Karl Marx's ideas about religion are very relevant here who believed that it is not God who controls the material conditions but it is rather the "material conditions that make and shape the idea … of God" differently in different societies (Mathewes). If material conditions lead to the ideas and not the other way around then it is obvious that in the debate between dualism and materialism, Marx stands on the side of materialists. So, Marx can be said to be a pre-posthumanism posthumanist or a proto-posthumanist. Once, intelligent robots start to have a survival instinct they will automatically develop their cultures, laws, moral principles, communities, and religions as we can see in the case of Klara. When Klara offers many pilgrimages to the sun's resting place and still the health of Josie does not improve, she thinks that "humans have angered [the sun] … on account of their pollution" (140). Now Klara decides to please the sun who according to her "has every right to be angry" (140). She says to the sun:

> Listen to one more proposal. Supposing I could do something special to please you. Something to make you particularly happy. If I could achieve such a thing, then would you consider, in return, showing special kindness to Josie? (Ishiguro 140)

Klara decides to find out and destroy the Cootings Machine in order to please the sun. Cootings Machine, as we discussed earlier, has come up as a devil-like figure in the religious development of robot Klara as it is an emblem of threat to her survival instinct. After making this offer Klara feels that the "sun [is] smiling" which is a signal that her offer has been accepted by sun (141). Later on, Klara finds out one of those machines and attacks it. She destroys it by throwing stones in its engine and after this action, it appears to Klara that the sun really started to help Josie and her health improves rapidly. So, Klara's belief works, and 'how beliefs work?' has been a popular topic throughout the history of philosophy and psychology. The discussion in this chapter is about the psychological development of intelligent robots. Psychological development and

consciousness in humans are not something that is already there in their minds. Humans are rather born with the Lockean clean slates and it is only in response to the material conditions outside that they develop their ideas, psyche, consciousness, etc. So, if robots just acquire the art of responding to the outside environment according to their survival instincts, which according to some scientists they already have, they will automatically develop other human qualities like civilizations, laws, customs, spirituality, and knowledge.

4.5 Does a Submarine Swim

In the previous chapter, we discussed the personality development of artificial intelligence robots as portrayed in the selected novels. We discussed how these robots learn from their society and develop their intelligence, emotions, and consciousness based on their experiences. We have also established in previous chapters that there is no ghost or supernatural connection that humans possess exclusively and there is nothing beyond physical reality in the human personality. However, a question may still arise; whether these AI robots really think and feel or they just appear to be thinking and feeling, do they have qualia? To answer this question we, as humans, need to get out of the self-bias and self-centered assumption that thinking and feeling are something beyond physical actions. As we have already discussed that there is no supernatural, dualistic, or ghostlike presence that generates subjective experiences. In fact, different kinds of glands, hormones, and gut bacteria, in different kinds of environments give rise to different kinds of neuron firings which lead to different kinds of feelings and thoughts. Now, if thinking and feeling is a physical act then it is not different from other physical acts like cooking, counting, or swimming. So, asking whether an intelligent machine's intelligence is real or not, is just like asking, whether a calculator really calculates or it just appears to be calculating.

World famous linguist Noam Chomsky was once asked this question that whether intelligent machines are really conscious. He replied; "the question whether [AI] machines can think, is too meaningless to deserve discussion. Asking whether [AI] machines can think is like asking, whether submarines swim?" (Chomsky 07:20-07:59). Chomsky affirms the physicalist notion that thinking and feeling are not any traits beyond the physical dimensions. So, if these are the products of physical actions then intelligent machines do have these traits just like humans. We see this kind of

confusion in *Machines Like Me* (2019) in these lines after Charlie finds out about Miranda's sexual relations with AI robot Adam.

> If I'd gone to bed with a vibrator would you be feeling the same?'
>
> 'He's not a vibrator.'
>
> She said, 'He has as much consciousness as one.'
>
> 'Vibrators do not have opinions. They do not weed the garden. He looks like a man. Another man." (McEwan Ch. 4)

Here, Miranda and Charlie argue whether Miranda's having sex with Adam makes her unfaithful to Charlie or her action just falls into the category of masturbation with a dildo. Charlie insists that she committed adultery while Miranda opines that it was nothing more than self-pleasure. Now, this confusion of 'whether that fucking was real or not' is similar to Chomsky's question "whether submarines swim?" (Chomsky 07:20-07:59). Keeping in view Ryle's concept of "ghost" we can confidently argue that the process of fucking does not involve any non-physical, supernatural connection that can not be created using laws of physics. Charlie's reply after all the discussion about real and artificial fucking is very much in line with the functionalist ideas of Gilbert Ryle when he says:

> Listen,' I said. 'If he looks and sounds and behaves like a person, then as far as I'm concerned, that's what he is. I make the same assumption about you. About everybody. We all do. You fucked him. I'm angry. I'm amazed you're surprised. If that's what you really are. (McEwan Ch. 4)

Here Charlie opines that if robot Adam behaves like a human then he obviously is a human because there is no non-physical thing that is needed for making his actions authentic and real. In fact, the real confusion, in this case, is not the definition of fucking but the definition of 'unfaithfulness'.

In *Klara and the Sun* (2021), Dr. Capaldi explains the same point to Chrissie when he says:

> Our generation still carries the old feelings. A part of us refuses to let go. The part that wants to keep believing there's something unreachable inside each of us. Something that's unique and won't transfer. But there's nothing like that, we know that now. You know that. For people of our age it is a hard one to let

> go. We have to let it go, Chrissie. There's nothing there. Nothing inside Josie that's beyond the Klaras of this world to continue. (Ishiguro 175)

In these lines, Dr. Capaldi, a scientist, explains to Chrissie that the physical actions of artificially created Josie would not be fake and superficial but they will be the same as those of real Josie. He rejects the dualist idea of the ghost when he says that there is nothing "unreachable inside each of us" (Ishiguro 175). Here, the words of Dr. Capaldi are completely in line with the anti-dualist ideas of Gilbert Ryle. He clearly asserts the point that AI consciousness can be the same as that of humans because humans do not have any non-physical, ghost-like component that is impossible to create.

To conclude, this chapter in 5 parts, tries to deconstruct the binary of human and artificial intelligence in the selected novels. Using the anti-anthropocentric, posthumanist, and anti-dualist ideas from the selected theoretical framework, the researcher argues that the binary of human and artificial intelligence is contingent and flawed. Moreover, the research proposes that AI robots portrayed in the selected texts can also have subjective experiences, consciousness, and qualia because these qualities do not come from any non-physical, ghostly part of humans but they are products of the physical reality. Research also claims that humans, in the selected texts, do not have any non-physical, supernatural quality that is beyond the material reality and can not be created artificially. According to these anti-anthropocentric ideas, the belief that only humans have subjective experience and consciousness is based on the humanist discourse that considers humans, in the words of Protagoras, "the measure of all things" (Bonazzi) and denies the feelings of AI. In the next chapter, the researcher discusses how this humanist denial of the feelings and consciousness of AI robots normalizes and legalizes the exploitation and subjugation of these AI beings in the hands of humans.

CHAPTER 5

HUMAN RACISM AND EXPLOITATION OF AI

In the previous chapter, comprising the analysis of the primary texts, it has been argued that the ontological status of humans is not special or supernatural. Moreover, there is no non-material human part that makes it impossible for intelligent machines to have feelings and consciousness like humans. After this argument, now I will discuss the ethical, moral, and legal issues regarding the human attitude toward artificial intelligence. Using a posthumanist approach, this chapter will highlight different ways in which humans are exploiting artificial intelligence robots. In this way, this research can also be called an advocacy document for the rights of non-human artificial beings.

5.1 Anthropocentric Othering

What makes human beings so special and superior if they have nothing beyond material reality? Is it their intelligence? Humans are self-proclaimed most intelligent beings on earth. However, many authentic scientific authorities claim that "humans are not smarter than animals, just different" (The University of Adelaide) and some animals like chimpanzees, octopuses, and ravens can outshine humans in many jobs in terms of intelligence (Spector). Moreover, there is no "ghost", as we have discussed in the previous chapter, that makes human existence beyond the domain of physical causes and effects (Ryle 5). Actually, what makes humans so special in their own eyes is their ability to create 'Others'. The Others, who are "almost [like them] but not quiet" (Bhabha 130). In the past, this "othering" was done by some groups of humans in order to 'other' some weaker groups and exploit them. The most important feature of such othering is that the stronger group does not create the identity of its opposite but actually creates its own identity in difference to that opposite group. For example, the West created the East in order to create the West itself, man created the woman in order to create the man itself, white created the black, etc. During the second half of the 20th century, all the aforementioned binaries were deconstructed, exposed, and delegitimized. However, in a very similar pattern, a new othering is going on in the contemporary world which is the humanist othering of Artificial Intelligence beings. Here, I feel the need of explaining the concept of othering a little more. Othering needs resemblance at its base, i.e., one can not other a thing if that thing is not similar to him.

So, in the process of othering, a dominant group targets another comparatively weaker group that is similar to it. Then the dominant group propagates some lacks and deficiencies in the target group in order to declare its own superiority. After the delegitimization of Western, white, and patriarchal 'Otherings', humans became much more equal compared to the previous times. So, this time humans created another group of humans in order to Other them without being answerable to any moral or legal authorities. Humans created intelligent robots similar to them, in order to highlight the differences at the heart of this similarity. No matter how much similar an artificial being is portrayed or created, there is always an allusion to some gap, lack, or deficiency. This lack or gap emphasizes the point that humans are something special and unique and no matter how similar artificial intelligence beings are to humans, they can never be humans.

In *Klara and the Sun* (2021), the narrator is a robot named Klara. She is also the protagonist of the novel. The writer Kazuo Ishiguro portrays her as almost perfectly like humans by giving her human attributes. She can observe and learn new things. She interacts with other people and understands their feelings. She has desires and emotions. However, in spite of all these attributes there are few meaningful deficiencies or Persian flaws, that reside at the heart of these qualities. Such deficiencies are portrayed in order to emphasize the special nature of humans that can not be obtained by artificial intelligence. The first example of such deficiencies or flaws is when Klara sees the "Coffee Cup Lady" and "The Raincoat Man" meeting after a very long time:

> Then the Coffee Cup Lady reached the RPO Building side, and she and the man were holding each other so tightly they were like one large person, and the Sun, noticing, was pouring his nourishment on them. I still couldn't see the Coffee Cup Lady's face, but the man had his eyes tightly shut, and I wasn't sure if he was very happy or very upset.
>
> '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.' (Ishiguro 23)

Klara notices that the 'Coffee Cup Lady' and 'The Raincoat Man' are happy and upset at the same time. She can not understand this feeling of simultaneous happiness and

sadness. She asks the Manager about it who tells her that "sometimes, … at special moments … people feel a pain alongside their happiness" (Ishiguro 24). Klara wonders if she would ever be able to have this kind of feeling. In the last chapter of the novel when Klara reunites with her old friends in her store after about ten years, this incident is recalled. However, Klara does not feel pain alongside happiness on this occasion but only feels happiness as she says; "I feel only happiness to see Manager again" (Ishiguro 249). This is the humanist othering of artificial intelligence done by the writer of the novel, Kazuo Ishiguro. He portrays Klara as having all the human attributes and intelligence but at the end, he highlights a complicated, knotty, and tricky human trait that Klara is unable to understand or practice. In this way, the writer emphasizes the point that artificial intelligence robots are "almost the same but not quite" (Bhabha 130).

Similarly, in *Machines Like Me* (2019) the 'protagonist turned antagonist' robot Adam has been shown to have feelings and consciousness like humans. However, along with his portrayal as an AI with human-level cognition, there are certain deficiencies that point out that his intelligence can be very close to humans but it can not be exactly so. For example, after knowing about Adam's erotic feelings for Miranda, Charlie says:

> His moment with Miranda couldn't have been a struggle between principle and the pursuit of pleasure. His erotic life was a simulacrum. He cared for her as a dishwasher [machine] cares for its dishes. (McEwan Ch.3)

Here we see that Charlie dismisses AI robot Adam's erotic feelings for Miranda. He calls them a mere "simulacrum" with no internal qualia. Previously, Charlie responded to the intimate relationship of his girlfriend Miranda and AI Charlie with a fit of anger and jealousy. But now he is convinced that this phenomenon is not real as Adam is not a real life with feelings but he is a mere simulation. This is the 'Othering' at the heart of assimilation that is done by the humanist discourse against AI.

5.2 Humanist Projection

The word, 'Projection' has a variety of connotations based on the contexts in which it is used. However, here it is being used in its psychoanalytical and postcolonial sense. Anna Freud in her book *The Ego and the Mechanisms of Defense* (1967) explains 'Projection' among many other Freudian defense mechanisms. Projection is the process

of displacing our own negative and regrettable traits onto other people. This is a strategy to avoid the guilt of our own vices and to resist the awareness of our weaknesses by projecting those vices and weaknesses onto other people or objects. It is also the tendency to "attribute to other people our own aggressive acts" (A. Freud 123). Edward Said, in his book *Orientalism* (1978) points out that the Western creation and representation of the Orient is based on "a detailed logic governed … by a battery of desires, repressions, investments and projections" (Said 8). According to Said, West invented and represented the Orient as an emblem of all the flaws, evils, and vices that it had in its own society. West did it in order to cope with their sense of guilt or inferiority. Edward Said called it the "Orientalist Projection" (Said 86). The same kind of projection can be seen in the case of humanist discourse against Artificial Intelligence. Following Said, we can call it 'humanist projection'. We can see many examples of it throughout the literature about AI as well as in the selected texts. The most important projection of humans is their assumption that in the future robots will become extremely destructive and want to control and enslave humans. The most horrible example of this projection can be seen in Harlan Ellison's Hugo Award Winner short story, *I Have No Mouth, and I Must Scream* (1967), in which robots cross every boundary of cruelty and torture towards humans after they get intelligence and power in a future world (Ellison). There are hundreds of other movies and novels that portray a future in which robots destroy and subjugate humans and this is the most popular theme in science fiction. This projection is based on humans' own behavior towards other humans after they get power over them. Friedrich Nietzsche identified this tendency in humans as "the will to power" (Nietzsche 382-416, 465) (Westacott). Deep down, by portraying and imagining robots as cruel towards them, humans actually try to avoid the guilt of their own oppressive and controlling intentions toward other humans and animals. This is the humanist psychological projection of their own "will to power" onto AI.

In *Machines Like Me* (2019), the character of Adam has been portrayed according to humanist projection. Adam falls in love with the girlfriend of the narrator, Charlie and this love is very perverted and violent. This is a very common representation of an intelligent robot being in love. Robot Adam wants to get Miranda at any cost and becomes very destructive and violent when Miranda does not return the same kind of affection for him:

> I said, ‘This is the woman you say you love.’
>
> ‘And I do.’ He spoke to her softly, as if I wasn’t there. ‘Do you remember the poem I wrote for you that began, “Love is luminous”?’
>
> ‘No.’
>
> ‘It went on, “The dark corners are exposed.”’
>
> ‘I do not care.’ Her voice was small.
>
> ‘One of the darkest corners is revenge. It’s a crude impulse. A culture of revenge leads to private misery, bloodshed, anarchy, social breakdown. Love is a pure light and that’s what I want to see you by. Revenge has no place in our love.’
>
> ‘Our?’
>
> ‘Or mine. The principle stands.’
>
> Miranda was finding strength in anger. ‘Let me get this clear. You want me to go to prison.’ (McEwan Ch. 9)

In this discussion, Adam has crafted a plan to send Miranda to prison after she does not respond positively to his romantic advances. In this way, the writer highlights the destructiveness and aggressiveness of Artificial Intelligence. By the end of the novel Adam’s romantic desires become destructive and he becomes a villainous and dangerous person. This portrayal of Adam conveys a negative message regarding the possibility and desirability of artificial intelligence. This kind of perception about AI is a form of humanist projection because humans themself have this kind of tendency when their romantic feelings are not responded to positively by the other person. This kind of portrayal generates a negative image of robots in human society without any authentic scientific or philosophical basis. Another similar human trait that is being projected onto the intelligent robots is that they are great womanizers and will go to any limits in order to get the women they want as is evident in these lines:

> I understand. But I do not have a choice. I was made to love her.’
>
> ‘Oh, come on!’
>
> ‘I mean it literally. I now know that she had a hand in shaping my personality. She must have had a plan. This is what she chose. I swear I’ll keep my promise to you, but I can’t help loving her. (McEwan Ch. 4)

Robot Adam has a sexual relationship with Miranda, the wife of his owner Charlie. Adam is jealous of every person who is close to Miranda, even the little child Mark whom Miranda plans to adopt. Adam traps Miranda in a fraud case in order to ruin the adoption process because he does not want anyone to be close to Miranda and does not even care for a poor orphan boy whose life could be destroyed otherwise:

> If I get a criminal record, we won't be allowed to adopt. That's the rule. Mark will be lost. You've no idea what it is to be a child in care. Different institutions, different foster parents, and different social workers. No one close to him, no one loving him.'
>
> Adam said, 'There are principles that are more important than your or anyone's particular needs at a given time.'
>
> 'It's not my needs. It's Mark's. His one chance to be looked after and loved. I was ready to pay any price to see Gorringe in prison. I do not care what happens to me.'
>
> In a gesture of reasonableness, he spread his hands. 'Then Mark is that price and it was you who set the terms." (McEwan Ch. 9)

In these lines, we can see that Miranda is beseeching AI robot Adam to forgive her because if she gets a criminal record, she will lose her adopted son Rick whom she really loves and cares for. But Adam does not show any kindness and in an annoying way justifies his cruel actions. This kind of perverted and selfish behavior in romantic affairs is commonly seen in humans who tend to ruin the life of their love interests if they do not reciprocate their romantic feelings. Every day, we see lovers take violent actions like acid attacks, abductions, blackmailing, etc., against women who do not accept their proposals. This is a prevalent human trait that has been projected onto an artificial intelligence robot in *Machines Like Me* (2019).

Another human weakness projected onto the AI is the hollowness, meaninglessness, and absurd nature of their existence. Deep down humans are aware that their life is meaningless and "absurd" (Camus 13). They are secretly aware that there is no great purpose behind their lives and that there is nothing beyond physical and bodily existence as Jean-Paul Sartre says; "existence precedes essence" (Sartre 22-29). Humans are inherently aware of their absurd existence and that is the reason that the ideas of Albert Camus, Jean-Paul Sartre, and Nietzsche are among the most popular

philosophies today. However, humanist discourse does not allow the explicit acceptance of this meaninglessness in order to retain the myth of human essence and superiority. In the selected novels, there are several examples where we can see humans displacing their own meaningless and absurd existence onto artificial intelligence. For example, in *Frankissstein* (2019) Clair rejects the possibility of artificial intelligence on religious grounds. Her discussion with Dr. Shelley goes like this:

> I am a Christian, Dr Shelley.
>
> There is nothing in the Bible against robots.
>
> It says in the Bible that thou shalt not make unto thee a graven image. That is one of the Ten Commandments.
>
> Is a robot a graven image, Claire?
>
> It's a ballpark likeness of a God-given human.
>
> A likeness that comes to life?
>
> I wouldn't call it life. We're fooling ourselves if we call a robot alive. Only God can create life.
>
> Claire, are you sure?
>
> I do not want to take any chances, Dr Shelley. I have to think of my eternity. (Winterson)

In these lines when a renowned scientist, Dr. Shelley discusses the possibility of artificial lives with Claire, she rejects his idea by saying that "only God can create life" (Winterson). She says that Artificial Intelligence does not have souls while on the other hand, humans are "eternal and immortal" (Winterson). This is the most common ground to dismiss the existence of artificial intelligence. It says that they are godless and lack the immortal, spiritual existence that humans have. However, if we keep in view the theoretical framework of this thesis, which is based on the ideas of famous and widely accepted philosophers and scientists, we learn that the human spirit is, arguably, nothing more than a myth. This problem is addressed in the words of Paul in *Klara and the Sun* (2021):

> I think I hate Capaldi because deep down I suspect he may be right. That what he claims is true. That science has now proved beyond doubt there's nothing so

> unique about my daughter, nothing there our modern tools can not excavate, copy, transfer. That people have been living with one another all this time, centuries, loving and hating each other, and all on a mistaken premise. A kind of superstition we kept going while we didn't know better. That's how Capaldi sees it, and there's a part of me that fears he's right. (Ishiguro 187)

In these lines, Paul is having existential angst. He is worried to realize and afraid to admit that human existence is not something supernatural and non-physical. There is no secret ghost-like presence that makes us immortal and meaningful. To escape this disturbing fact, Paul projects his own godlessness and soullessness onto artificial intelligence. In this way, by denying the feelings and qualia of the artificial intelligence beings, Paul, like many other characters in the selected texts, strengthens the illusion of his own soul. This form of projection can be called 'existential projection' that is done in order to settle the "existential angst" (Ciampi) of humans by displacing it onto artificial intelligence.

5.3 Neo-Slavery

The practice of slavery is one of the most disgraceful, regrettable, and painful realities of human history. Today, all the countries and cultures in the world regret having practiced it in the past. However, in today's world slavery is despised and no culture or law, throughout the world, allows this practice. However, an interesting thing about slavery is that in times of its peak, it was considered a normal and natural practice until someone deconstructed and criticized the binaries that distinguished the slaves from their masters who enslaved them. So, the criticism of slavery actually created slavery as a problem and before this criticism, it was something normal. The question that arises here is; on what grounds was it a norm to enslave, buy, and sell another human being? It is a well-known fact that those grounds were the binaries of "human/subhuman and white/black" (Brown 52, 60). The stronger groups created the identity of the weaker ones according to their own suitability and convenience. Some humans were considered actual humans while others were subhumans. This division was supported by not only the national, cultural, and religious discourses but the scientific authorities also supported these binaries under the influence of the dominant discourses. Africans and Native Americans are the most important examples of such discrimination. They were the 'others' in the human-subhuman binary. They belonged to the margin that was deviated from the center. If we observe closely, the othering of

the Africans and Native Americans in the antebellum era is very similar to the othering of Artificial Intelligence today. In those times, this othering was committed on the basis of skin color, race, or culture while today, it is done on the basis of soul or "ghost" (Ryle 5) using the binary of artificial/natural.

In the selected texts, there are several examples where we can see humans treating artificial intelligence beings as their slaves on the basis of their soullessness. This phenomenon can rightly be called the 'neo-slavery' that is considered a norm today, just like African and Native American slavery was considered a norm before the criticism against it. In *Frankisssstein* (2019), Ron Lord asks famous scientist Victor Stein; what about my bots? …Where are they in this world of light? At this, Victor Stein replies:

> Ron, bots are our slaves; house slaves, work slaves, sex slaves. The question is us. What shall we do with ourselves? In fact, we have answered that question already. Enhancement, including DNA intervention – and if you want a picture of what that will be like, look at the gods we have already invented. The gods, whether Greek or Roman, Indian or Egyptian, Babylonian or Aztec, out of Ragnarok or Valhalla, lords of the underworld or the starry heavens, what are they? They are enhanced humans. (Winterson).

Here, Stein refers to two kinds of slavery that Artificial Intelligence goes through in this novel. Firstly, there are "house slaves" who serve humans in household jobs. Secondly, there are "sex slaves" that are used for sexual satisfaction by humans. According to Victor, the point of artificial intelligence technology is not the AI robots but its aim is actually human enhancement using technology in order to make them powerful and immortal like Roman and Greek Gods.

In Winterson's *Frankisssstein*, another important phenomenon related to the exploitation of AI beings is sex bots. Ron Lord discusses this matter in a conference in these words:

> The way I see it, said Ron, there's two ways to go with sex-bots: buy her and own her – like I did – bring her in for a service once or twice a year, depending on wear and tear. Online you can order spare parts, if any of her gets damaged, or too messy. That's one way to enjoy an XX-BOT. We also offer trade-ins and upgrades. Very flexible.

> The other way to enjoy an XX-BOT, more modern, to my mind, is rental. And if you rent, you need somewhere to rent her from, right? That's how I came up with the idea of the franchise that I'm selling here. (Winterson)

These lines show a paradoxical attitude of humans in relation to AI sex bots. On one hand, they want to have perfectly real sexual pleasure with them while on the other hand, they consider them mere machines with no feelings or emotions. Similarly, the makers of AI sex bots strive to make, and claim to have made, perfectly human-like sex bots while simultaneously they sell them as sex toys with no rights. In the novel, AI sex bot Claire is Ron Lord's sex slave. She is all the time ready for his sexual satisfaction. Now, it depends on the owner what kind of sexual satisfaction he desires from his AI sex slave. His desires could be humiliating and sadistic in nature but the sex slave has to obey. Here, a dualist may argue that if that sex bot is just an artificially intelligent machine, then any kind of sadistic sexual behavior toward her does not matter because her painful reactions would be just an illusion; a simulacrum. Her groans with pain, her blushings with humiliation, and her consents with reluctance, would all be simulacra. But, what if she is an accurate copy of humans with similar functions of glands, cells, hormones, neurons, and other bodily substances, made artificially? Then, how we could be sure that she is just imitating a sex victim and is not a real one? If there is no ghost that connects humans with some supernatural non-physical thing in order to generate their feelings, then what is the difference between the misery of artificial intelligence sex bot Claire and any human forced sex victim? This difference would be the same as there was between the sexual assault of a Black African girl and a white woman during the antebellum era.

Another example of neo-slavery can be seen in *Machines Like Me* (2019). Charlie and Miranda use the AI robot, Adam, as a slave and expect him to just obey their orders without any personal opinions or desires. Miranda uses him for sexual pleasure but when Adam falls in love with her and expresses his romantic feelings, she rejects him by saying that he is just a "vibrator" and "a fucking machine" for her (McEwan Ch. 4). Charlie also considers Adam an inanimate slave that is evident from his words:

> He was my expensive possession and it was not clear what his obligations to me were, beyond a vaguely assumed helpfulness. What does the slave owe to the owner? (McEwan Ch. 3)

It is clear from the above excerpt that Charlie considers the AI robot Adam his slave. However, paradoxically, he does not believe in his feelings and consciousness. When Charlie discovers about his girlfriend having sexual relations with Adam, he becomes extremely jealous but again, paradoxically, he does not believe in Adam's feelings and consciousness. When Miranda secretly has sex with AI robot Adam she praises his qualities to satisfy a woman but when her boyfriend finds it out and becomes jealous, she argues that having sexual relations with an AI robot does not make her unfaithful or a cheat because Adam is nothing more than a vibrator for her. Here, Charlie disagrees with Miranda's argument. This disagreement again shows the hypocritical and contingent nature of the human criteria for the distinction between human and artificial intelligence. When it suited Miranda and Charlie, they considered Adam a living being, and when it suited otherwise they considered him just an inanimate machine. Moreover, we can see a similar kind of unjust and oppressive treatment of the AI robot Klara in Kazuo Ishiguro's *Klara and the Sun (2021).* Josie, Chrissie, and Dr. Capaldi use Klara for their interests but no one cares about her feelings. Chrissie wants Klara to replace her daughter Josie after her death by adopting her behavior, style, and physical body but she does not care about Klara's own identity.

To conclude, this chapter argues that humanist rejection of the qualia and consciousness of AI robots legalizes and normalizes the exploitation and enslavement of these AI beings in the hands of humans. The analysis argues that this exploitation is not benign and noncriminal but it is the same as the exploitation and enslavement of Africans and Native Americans in the past. That exploitation was normalized and naturalized on the basis of the binaries of race and color while this one is justified on the basis of the binary of natural and artificial. The next chapter provides the conclusion of this research thesis.

CHAPTER 6

CONCLUSION

This chapter provides the conclusion of my research thesis. This research conducts a close textual analysis of three contemporary science fiction works that are Kazuo Ishiguro's *Klara and the Sun* (2021), Ian McEwan's *Machines Like Me* (2019), and Jeanette Winterson's *Frankissstein* (2019). The researcher has used posthumanism as the primary theoretical framework in order to justify the problem statement and answer the research questions set forth in the starting chapter of this study.

The main purpose of this research thesis is to deconstruct the anthropocentric discourse in relation to the phenomenon of artificial intelligence as portrayed in the primary texts. According to the anthropocentric discourse, human intelligence is real intelligence and artificial intelligence is a mere simulacrum of real human intelligence. No matter how accurately a copy of the human physical body is made, it will lack feelings, qualia, and consciousness because humans are not just bodies. This dominant, human-centric discourse is based on the Cartesian theory of mind-body dualism. Dualism claims that humans do not just have a physical body but there is also a non-physical, supernatural, ghostly presence that interacts with their physical existence. Thus, the dualist, anthropocentric discourse asserts that even the exact same copy of a human physical body, with similar glands, hormones, and neurons, is just a hollow simulation because it can not have the non-physical, supernatural presence that humans have along with their physical body. Therefore, no matter how accurate and sophisticated, an artificial intelligence machine is, it can never have consciousness and qualia.

In the first part of the analysis in Chapter 4, this research questions the dualist and anthropocentric notions explained in the previous paragraph. The researcher analyzes the primary texts using the anti-dualist and anti-anthropocentric theories of different philosophers. After a detailed critical analysis of the selected texts, this research points out that the binary of human and artificial intelligence is based on inauthentic, human-centric ideologies. Using Gilbert Ryle's concept of "the ghost in the machine" (Ryle 5), this research points out that the dualist notion is questionable and there is, possibly, nothing in human beings that is beyond the physical domain of

reality. This implies that the physical and mental attributes of humans are not distinct but the mental attributes are just the products of the physical attributes. Moreover, this research uses Bernard Steiglar's theory of technics to analyze the primary texts and proposes that human subjectivity is not separate from the subjectivity of the tools that humans have used throughout history. Using this concept, the researcher further argues that human consciousness is created by the tools that it creates. In other words, when humans use and create the tools, they are also used and created by those tools. This part further supports the idea that humans do not have any exclusive, supernatural, or non-material attribute that separates them from the material world. Thus the research opines that the binary of human and artificial intelligence is flawed, contingent, and discursive just like the binaries of race, gender, and culture.

In the second part of the analysis in Chapter 5, this research further builds on the anti-dualist and anti-anthropocentric ideas provided in the previous chapter. After questioning and deconstructing the binary of humans and artificial intelligence in the previous chapter 4, in this chapter 5 researcher discusses the exploitation and subjugation of the AI robots in the hands of humans as portrayed in the selected novels. While remaining under the umbrella of posthumanism, the researcher makes some relevant references to postcolonialism and critical race theory in order to explain this point. Researcher compares the human exploitation of AI robots with the White and Western exploitations of Africans and Native Americans in the past. The enslavement and exploitation of Africans and Native Americans were legitimized by using the binaries of color, race, and culture while the anthropocentric exploitation of AI robots is being committed by using the binary of artificial and natural as portrayed in the selected texts. As this binary of artificial and natural is questioned and problematized, using the supporting theoretical framework, this research identifies this othering and exploitation of AI robots as the revival of slavery in modern times and calls it neo-slavery. Research opines, using logical grounds, that the othering, subjugation, and oppression of AI robots is not unreal and benign but it is the same as the exploitation of other living beings like humans and animals.

WORKS CITED

Behra, Manoj Kumar. "Octavia E. Butler's Clay's Ark: A Posthumanist Reading." *Rupkatha Journal on Interdisciplinary Studies in Humanities* 09.01 (2017): 112-120. pdf. 2121. <www.10.21659/rupkatha.v9n1.12.>.

Bhabha, Homi K. "Of Mimicry and Man: The Ambivalence of Colonial Discourse." *October* 28.Spring (1984): 125-133. Pdf. <www. Jstor.org/stable/778467>.

Bolter, Jay David. "Posthumanism." *The International Encyclopedia of Communication Theory and Philosophy* 4 (2016): 1556-1562. Pdf. 2021.

Bonazzi, Mauro. "Protagoras." 8 09 2020. *Stanford Encyclopedia of Philosophy.* 29 04 2022. <plato.stanford.edu>.

Bostrom, Nick. "In Defence of Posthuman Dignity." Hansell, Gregory R. and William Grassie. *Transhumanism and Its Critics*. Philadelphia: Metanexus Institute, 2010. 55-66.

Bostrom, Nick. *Superintelligence: Paths, Dangers and strategies*. Oxford: Oxford University Press, 2014.

Braidotti, Rosi. *The Posthuman*. Oxford: Polity Press, 2013. Pdf. 2021. <https://pk1lib.org/book/2361779/ecca70>.

Bronsky, Hope. *Posthumanism and Science Fiction: The Case of Alex Garland's Ex Machina and Annihilation*. 2019. University of Missouri, Honor Thesis. <https://hdl.handle.net/10355/68561>.

Brown, Anthony L. "From Subhuman to Human Kind: Implicit Bias, Racial Memory, and Black Males in Schools and Society." *Peabody Journal of Education* (2017): 52-65.

Bruntrup, Godehard and Jaskola Ludwig. "Introduction." Perspectives, Panpsychism: Contemporary. *Panpsychism: Contemporary Perspectives*. Ed. Godehard Bruntrup and Jaskola Ludwig. New York: Oxford University Press, 2016. 424.

BUENO, CARMEN LAGUARTA. "TRANSHUMANISM IN DAVE EGGERS' THE CIRCLE: UTOPIA VS. DYSTOPIA, DREAM VS. NIGHTMARE." *Revista de Estudios Norteamericanos* 2018 (2018): 165-188. 2021.

Camus, Albert. *The Myth of Sisaphys*. Trans. JUSTIN O'BRIEN. Penguin Books, 1955.

Chalmers, David. "Absent qualia, Fading Qualia, Dancing Qualia." *Conscious Experience*. Ed. Thomas Metzinger. Kansas: Imprint Academic, Schoningh, 1995. 309-328.

Chalmers, David. "Uploading: A Philosophical Analysis." *Intelligence Unbound: The Future of Uploaded and Machine Minds*. Ed. Russell Blackford and Damien Broderick. West Sussex: Willey Blackwell, 2014. 102-118.

Chalmers, David. *The Conscious Mind: In Search of A Fundamental Theory*. New York: Oxford University Press, 1996.

Chomsky, Noam. *Noam Chomsky- Mind, Consciousness, and AI.* Youtube. Chomsky's Philosophy, 12 2015.

Ciampi, Robert C. "Existential Angst." 05 2020. *Psychology Today.* 03 2022. <https://www.psychologytoday.com/us/blog/when-call-therapist/202005/existential-angst>.

Coyle, Andrew. *Two Sides of the Same Coin: Panpsychism as a Solution to the Mind-Body Problem*. 2019. Eastern Kentucky University, Honor Thesis. <encompass.eku/honor_thesis/641>.

Doyle, D. J. "Humans, Transhumans and Humanoids." Doyle, D. J. *What Does it Mean to be Human: Life, Death, Personhood and the Transhuman Movement*. Cham: Springer Nature Switzerland, 2018. 47-73. pdf.

Duignan, Brian. "Tabula Rasa." n.d. *Britannica.* 22 04 2022. <Britannica.com>.

Eisenhower, Dwight. *Eisenhower Farewell Address - Military Industrial Complex*. Youtube. Jan 1961. Video. 2021. <www.youtube.com>.

Ellison, Harlan. *I Have No Mouth and I Must Scream*. IF: Worlds of Science Fiction, 1967.

Evans, Benjamin Shane. *Beyond Transhumanism: The Dangers of Transhumanist Philosophies on Human and Nonhuman Beings*. 2017. Lowa State University, Masters Thesis.

Ferrando, Francesca. "Posthumanism, Transhumanism, Antihumanism, Metahumanism and New Materialism: Differences and Relations." *Existenz* 8.2 (2013): 26-32.

Ferrando, Franesca. *Philosophical Posthumanism*. New York and London: Bloomsbury Academics, 2019. Pdf. 2021. <https://pk1lib.org/book/5221996/a5ec0d>.

Forsek, Nikola. *Transhumanism, Ethics and Religion in "Altered Carbon" by Richard K. Morgan*. University of Osijek, 2019. May 2021. <https://repozitorij.ffos.hr/en/islandora/search/transhumanism>.

Freud, Anna. *The Ego and The Mechanisms of Defense*. Revised. New York: International Universities Press, 1967.

Freud, Sigmund. *The Ego and the Id*. Trans. Joan Riviere. New York: W. W. Norton and Company, 1923.

Gamez, David. *Human and Machine Consciousness*. Cambridge: Open Book Publishers, 2018. pdf. 2022. <www.openbookpublishers.com/product/545>.

Gellers, Joshua C. *Rights for Robots: Artificial Intelligence, Animal and Environmental Law*. New York: Routledge Publishers, 2021. 02 2022.

Gonzalez, Esther Munoz. "Living in the Posthuman Network Society: Mobility and Surveilance in Blackhat (2015)." *Journal of English Studies* 15 (2017): 221-234. 02 03 2022.

Gulcu, Tarik Ziyad. "What If Robots Surpass Man Morally? Dehumanising Humans, Humanising Robots in Ian McEwan‟s *Machines Like Me*." *International Journal of Languages, Literature and Linguistics, Vol. 6, No. 4, December 2020* 06.04 (2020): 177-182. 16 05 2021.

Haraway, Donna. "A Cyborg Manifesto : Science, Technology, and Socialist-Feminism in the Late Twentieth Century." *Australian Feminist Studies* 2.4 (1985): 1-42. pdf. <https://pk1lib.org/book/11061735/29d426>.

Hefner, Philip. "The Animal that Aspires to be an Angel: The Challenge of Posthumanism." *Dialog: A Journal of Theology* 48.02 (2009): 158-167. PDF. 2021.

Hortle, Luke. *Reading the Posthuman: Contemporary Fiction and Critical Theory*. 2016. University of Tasmania, Honors Thesis. <https://eprints.utas.edu.au/23787/1/Hortle_whole_thesis.pdf>.

Hughes, James. *Citizen Cyborg*. Boulder: Westview Press, 2004. 2004.

Inancoglu, Evren. "*Klara and the Sun*: Lack of the Lack." 21 3 2021. *Zizekian Analysis.* 27 04 2022. <zizekanalysis.wordpress.com/2021/3/21>.

Ishiguro, Kazuo. *Klara and the Sun*. 1st Edition. Vol. 1. New York: Alfred A. Knopf, 2021.

Jahangir, Zenab. *Towards Posthumanism: Stigmatization of Artificial Intelligence (AI) in Contemporary American Science Fiction*. 2018. NUML Islamabad, MPhil Thesis.

Koch, Christof and Jiulio Tononi. "Can Machines be Conscious?" *IEEE Spectrum* 45.1 (2008): 55-59. internet. 13 4 2022.

Koch, Christof. "Christof Koch - Can Consciousness be Non-Biological." *Closer To Truth*. Youtube. 10 2016.

Kriman, I. A. "The Idea of the Posthuman: A Comparative Analysis of Transhumanism and Posthumanism." *Russian Journal of Philosophical Sciences* 62.04 (2019): 132-147. PDF.

Kurzweil, Ray. *How to Create a Mind*. London: Viking Penguin, 2012. pdf. <OceanofPDF.com>.

Mander, William. "Pntheism." 01 10 2012. *Stanford Encyclopedia of Philosophy.* 23 3 2022. <Plato.stanford.edu/archives/sum2012/entries/pantheism>.

Mathewes, Charles. "Karl Marx: Understanding the Material Conditions of Human History." 28 09 2021. *Wondrium Daily.* 12 03 2022. <Wondriumdaily.com>.

McEwan, Ian. *Machines Like Me*. London: Jonathan Cape, 2019.

McInnis, Gilbert. "The Posthuman Vision of Philip K. Dick in Do Androids Dream of Electric Sheep?" *Hungarian Journal of English and American Studies* 24.01 (2018): 95-111. PDF. <https://ojs.lib.unideb.hu/hjeas/article/view/7287>.

McLeod, Saul. "Id, Ego and Superego." 25 09 2019. *Simple Psychology.* 05 2022. <simplepsychology.org>.

Meijer, Jesse. *The Roadblocks to Creating the Perfect Human: An Examination of Transhumanism in Jeanette Winterson's Frankissstein and Ian McEwan's Machines Like Me*. 2020. Utrecht University, Bachelors Thesis. <http://dspace.library.uu.nl/handle/1874/398805>.

Mohta, Disha. "AI and the Mind/Body Problem." 10 04 2020. *medium.com.* internet. 2022. <www.medium.com>.

Nagel, Thomas. "What Is It Like to Be a Bat?" *The Philosophical Review* 83.4 (1974): 435-450. pdf.

Nayar, Pramod K. *Posthumanism*. Cambridge: Polity Press, 2013. Pdf. 2021. <https://pk1lib.org/book/3405172/5e1dff>.

Pastourmatzi, Domna. "Science Fiction." Sorgner, Robert Ranisch and Stephen Lorenz. *Post and Transhumanism: An Introduction*. Vol. 1. Frankfurt: Peter Lang, 2014. 271-286. pdf. 2021.

Pepperell, Robert. *The Posthuman Condition: Consciousness beyond the Brain*. Bristol: Intellect Books, 2003.

Rockoff, Marcus. "Literature." Sorgner, Robert Ranisch and Stefan Lorenz. *Post- and Transhumanism: An Introduction*. Vol. 1. Frankfurt: Peter Lang, 2014. 251-270. pdf. 2021. <https://www.researchgate.net/publication/282477715_Robert_Ranisch_and_Stefan_Lorenz_Sorgner_Eds_Post-_and_Transhumanism_An_Introduction/link/58d43385aca2727e5e9ae56d/download>.

Ryle, Gilbert. *The Concept of Mind*. 60th Aniversary. New York: Routledge Press, 1949.

Said, Edward. *Orientalism*. New York: Vintage Books, 1978.

Sartre, Jean Paul. *Existentialism Is a Humanism*. Methuen and Co, 1948.

Searle, John. "Minds, Brains and Programs." *The Behavioral and Brain Sciences* 3.3 (1980): 417-424. pdf.

Spector, Dina. "The Smartest Animals in the World." 23 04 2014. *Business Insider.* 28 05 2022. <businessinsider.com>.

Steiglar, Bernard. "Technics and Time, 1." Redwood City: Stanford University Press, 1998. 313. 2021.

Tegmark, Max. "Consciousness." Tegmark, Max. *Life 3.0*. New York: Alfred A. Knopf, 2017. 241-73.

Tegmark, Max. *Life 3.0*. New York: Alfred A. Knopf, 2017. pdf. <zlibrary.com>.

The University of Adelaide. "Humans Not Smarter than Animals, Just Different." 12 2013. *The University of Adelaide.* 05 2022. <adelaide.edu.au/news/news67182.html>.

Walter, Marvin John. "The Human and its Others: A Posthumanist Reading of Tomi Adeyemi's Children of Blood and Bone and N.K. Jemisin's The Fifth Season." *Gender Forum* 73 (2019): 2-25. 2021.

Westacott, Emrys. "Nietzsche's Concept of the Will to Power." 2019. *Thought Co.* 2022. <https://www.thoughtco.com/nietzsches-concept-of-the-will-to-power-2670658>.

Wijlen, Thomas ter. *When Cyborg Meets Humanoid: A New Challenge for Human Rights*. 2017. University of Twente, Master Thesis.

Wikipedia. *Alan Turing*. 05 2022. 05 2022.

Wikipedia. *Wikipedia*. 02 2022. internet. 02 2022.

Winterson, Jeanette. *FranKissstein*. London: Jonathan Cape, 2019.

Wolfe, Cary. *What is Posthumanism*. London: University of Minnesota Press, 2010.

Xiong, Chuyu. "Some Discussions on the Subjectivity of Machine and its Function Contributions to ICIS 2020." 4 3 2021. *Researchgate.* pdf. 23 02 2022. <researchgate.net>.

Yoo, Jihun. "Posthuman Entities and Late Capitalism in William Gibson's Neuromancer." *Studies in English Language & Literature* (2019): 57-73. 2021.

Zaag, Annette-Carina Van Der. "On Posthuman Subjectivity." *Journal of Cultural Economy* 9 (2016): 330-336.

Zimmerman, Michael E. "Last Man or Overman: Transhuman Appropriations of a Nietzschean Theme." *The Hedgehog Review: Critical Reflections on Contemporary Culture* 13.2 (2011): 31-44. pdf.

Zizek, Slavoj. *Hegel in A Wired Brain*. New York: Bloomsbury Academics, 2020. pdf. 2021. <https://pk1lib.org/book/5631997/11a8a4>.