



ARTIFICIAL INTELLIGENCE THROUGH
THE (RED) LOOKING GLASS
HUMANKIND'S FEARS AS MENTAL PROJECTIONS

JESPER HANSEN

AALBORG UNIVERSITET

2020

ARTIFICIAL INTELLIGENCE THROUGH THE (RED) LOOKING GLASS

Humankind's Fears as Mental Projections

Jesper Hansen

Engelsk Almen, Aalborg Universitet

Jørgen Riber Christensen

June 2nd, 2020

123.621 Keystrokes

ABSTRACT

Dette speciale har til formål at undersøge hvorfor kunstig intelligens oftest bliver portrætteret som fjenden i et narrativ. Dernæst vil det blive undersøgt hvorfor vi som mennesker finder tanken om kunstigt intelligente computere og robotter skræmmende. Til at undersøge dette vil der blive gjort brug af Murray Smiths teori om sympatistrukturen, for at undersøge om vi som seere kan danne sympati for en kunstig skabning, eller om det sker gennem manipulerende narrative redskaber. Ydermere vil Michael Szollosys teori om projektion af selvet blive benyttet til at undersøge hvad det er ved disse robotter der gør os utilpasse. Til sidst vil der også, i det omfang det er muligt, blive gjort brug af Freuds teori om 'The Uncanny', Cynthia A. Freeland's teori om det sublime, samt Stephen T. Asmas teori omkring robotter i forhold til monstrologi. Ved at analysere filmene *2001: A Space Odyssey*, *Ex Machina* og *Bicentennial Man*.

Gennem analysen er det blevet demonstreret at det er muligt at danne sympati for en kunstig skabning. Dette kan ske både gennem pålidelig information om en karakter, men denne sympati kan også blive brugt til at manipulere seeren til at sænke sine parader og derved efterlade seeren med en følelse af at være blevet forrådt. Ydermere har analysen påvist at robotter gør os utrygge fordi de kan minde os om vores egen visse død. Til sidst har dette speciale også påvist at vores frygt for robotter stammer fra vores egne voldelige projiceringer som robotten påtager sig og derved bliver et symbol på det, der er ondskabsfuldt og farligt.

TABLE OF CONTENTS

Introduction	1
Literature Review	2
Science Fiction and Philosophy: From Time Travel to Superintelligence.....	2
On Monsters: An Unnatural History of Our Worst Fears	4
Feeling Robots and Human Zombies: Mind Perception and The Uncanny Valley	5
Freud, Frankenstein and our Fear of Robots: Projection in our Cultural Perception of Technology.....	6
The ‘Uncanny’ (Das Unheimliche)	8
Film Selection	9
2001: A Space Odyssey (1968).....	9
Ex Machina (2014).....	11
Bicentennial Man (1999).....	11
Theory	13
Importance of Character	13
Character Engagement	15
The Structure of Sympathy	16
Recognition	16
Alignment.....	17
Allegiance	18
Frankenstein’s Monster Version 2.0	20

Projections	22
The Sublime	24
Analysis	28
2001: A Space Odyssey	28
The Sublime in 2001: A Space Odyssey	29
The Monolith as A Sublime Object.....	35
The Structure of Sympathy in 2001: A Space Odyssey	37
The Primates.....	38
Dr Haywood R. Floyd	39
Dr Frank Poole	41
Dr David ‘Dave’ Bowman	42
Hal	45
Hal as a Digital Monster.....	47
The Uncanny in 2001: A Space Odyssey	49
Ex Machina	50
The Structure of Sympathy in Ex Machina.....	50
Caleb Smith.....	51
Nathan Bateman	52
Ava	54
Making of a Beautiful Monster	56
The Uncanny in Ex Machina.....	58
Ava as A Sublime Object	59

Bicentennial Man.....	59
The Structure of Sympathy	60
Andrew	60
Monsters of Bicentennial Man	62
The Uncanny in Bicentennial Man.....	63
Discussing Artificial Intelligence: Friend, Foe or Slave?.....	65
Conclusion.....	67
References	69
Litterature	69
Film	71

INTRODUCTION

“I’m sorry, Dave. I’m afraid I can’t do that”. By reading that sentence, one’s mind can almost hear the HAL9000 computer’s bone chillingly calm voice. HAL9000, or simply Hal, has become the posterchild of artificially intelligent computers that turns evil and kill their masters. Since then, the film industry has shown us an array of artificially intelligent robots, who, like Hal, seeks out to kill the human race. A few examples being: Arnold Schwarzenegger’s T-800 in *The Terminator* (1984), Ultron in *Avengers: Age of Ultron* (2015), The rebelling replicants in *Blade Runner* (1982).

If we had to take something away from these films, it should be that one can never trust an artificially intelligent robot. Furthermore, when various renowned scientists and inventors such as Stephen Hawking and Elon Musk speak up, warning us about artificial intelligence, one might argue that we ought to listen (Sparkes 2013). Stephen Hawking even suggested that the rise of artificial intelligence could see the human race become extinct (2013).

However, it seems that Hollywood keeps making films about artificial intelligence rising up against their human makers, causing dread, terror, and destruction. And yet, we do not bat an eye at the artificial intelligent assistant in the pocket of our jeans, whether it is named Alexa, Siri, Bixby, or Google.

Thus, leading to the research question of this thesis:

Is it truly the artificially intelligent robots that we fear? Or is it the thought that we are able to conjure our own monsters out of violent fantasies, and cold unfeeling rationalism?

LITERATURE REVIEW

Firstly, this thesis will engage in a literature review to gain a comprehensive view of a number of selected books and articles that revolves the subject of artificial intelligence and robots. What is it that makes people feel uneasy or even scared of the idea of an artificially intelligent robot? The following section will create an outline of some of the selected texts that will lay the groundwork for the theory section of this project. In the literature review, the following texts will be examined and summarised:

- Asma, S. (2009). *On Monsters: An Unnatural History of Our Worst Fears*. Oxford; Oxford University Press.
- Freud, S., Strachey, J., Freud, A., & Richards, A. (1974). *The standard edition of the complete psychological works of Sigmund Freud*. (pp. 219-252) London: Hogarth Press and the Institute of Psycho-Analysis.
- Gray, K., & Wegner, D. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1), 125-130.
<https://doi.org/10.1016/j.cognition.2012.06.007>
- Schneider, S. (2016). Asimov's "Three Laws of Robotics" And Machine Metaethics. *Science fiction and philosophy: from time travel to superintelligence* (2nd ed.). (pp. 290-308) West Sussex, England: Wiley Blackwell
- Szollosy, M. (2017). *Freud, Frankenstein and our fear of robots: projection in our cultural perception of technology*. *AI & SOCIETY*, 32. (pp. 433–439.)
<https://doi.org/10.1007/s00146-016-0654-7>

SCIENCE FICTION AND PHILOSOPHY: FROM TIME TRAVEL TO SUPERINTELLIGENCE

Susan Leigh Anderson states that *Bicentennial Man* (Asimov 1984) lays the groundwork for Machine Metaethics since it discusses Asimov's "Three Laws of Robotics", which is the ethical code for intelligent robots. However, Anderson argues that the Asimov rejected his own Three Laws of Robotics with his novel *The Bicentennial Man* (Asimov 1984), later adapted for the screen as *Bicentennial Man*. In both the novel and the film, the robotic protagonist, Andrew,

possesses characteristics that should not make him a slave to human beings as the Three Laws dictate, both texts provide an explanation to why humans feel the need to treat intelligent robots as slaves, which exposes a weakness. A weakness which makes it difficult for them to act as ethical paragons (Schneider 2016). Because of this weakness, Andrew could, according to Anderson, be more ethical than most human beings.

Anderson utilises *The Bicentennial Man* as a springboard for a discussion of Machine Metaethics, which leads to the following conclusions: 1: A machine could follow ethical principles better than most human beings. 2: Developing a program that enables a machine to give human beings ethical advice is the first step in the Machine Ethics project. 3: An intelligent robot, like Andrew, satisfies most, if not all, of the requirements philosophers have proposed for an entity to have moral standings/rights, hereby, making the Three Laws immoral. 4: Even if the machines that are actually developed fail to be like Andrew and should likely not be considered to have moral standing/rights, it is still problematic for humans to program them to follow the Three Laws. Based on these conclusions, Anderson argues that Asimov's Three Laws of Robotics will be an unsatisfactory basis for Machine Ethics.

While *Bicentennial Man* raises some questions about moral and ethics between human and machine, it is also a reminder to the Americans about their history and how particular groups, African Americans, had to fight for their freedom to be accepted by other human beings. It was wrong to treat other people as slaves, which is essentially the idea with robotic servants as Andrew. Like the African Americans, the intelligent robots in *Bicentennial Man* are feared by humans.

Machine Metaethics examines and talks about the field of Machine ethics rather than doing work in it. It raises questions such as; what is the ultimate goal of Machine Ethics? What does it mean to add an ethical dimension to Machines? Is ethics computable? Is there a single correct ethical theory that we should try to implement? Should we expect the ethical

theory we implement to be complete, that is, should we expect it to tell the machine how to act in an ethical dilemma in which it might find itself? is it necessary to determine the moral status of the machine itself, it is to follow ethical principles? (Schneider 2016).

Throughout the paper, Anderson discusses ethical principles and dilemmas and what characteristic(s) that are necessary to have moral standing. Ultimately, she argues that the Three Laws of Robotics are unsatisfactory even if machines do not have a moral standing. The best starting point of making ethics computable is by creating a program that enables a machine to act as an ethical advisor to human beings. By doing so, it will not require that we make a judgement about the ethical status of the machine itself, a judgement that, according to Anderson, will be a difficult one to make. Therefore, Asimov's Three Laws of Robotics are an unsatisfactory basis for Machine Ethics, regardless of the status of the machine (Schneider 2016).

ON MONSTERS: AN UNNATURAL HISTORY OF OUR WORST FEARS

In the chapter *Future Monsters: Robots, Mutants and Posthuman Cyborgs*, Stephen T. Asma discusses the concept of monster contains a Masochistic dimension, which can be seen in *Frankenstein* and in the Christian and pagan traditions. This type of monster often arises to take revenge on deserving individuals and even cultures as a whole (Asma 2009). Asma retains that the religious or cosmic aspect of monsters is still retained in the monsterology; Godzilla is the human race's own fault in the same sense that "the creature" was Dr Frankenstein's fault. According to Freud, the technology can be seen as the human race's prosthetic attempt to become God, which by the Christian logic will result in human punishment in some kind of monster. This leads us to the creation of robots, and artificial intelligence which has created a new potential monster in the sense that Hollywood movies such as *Blade Runner* (1982) in which the robots (replicants) are used as slaves with a limited lifespan to avoid a robotic rebel,

which ultimately will lead to the human race's demise. Such Hollywood narratives have raised all kinds of interesting questions about free will and the evolution of self-agency in a digital system. When does a machine become a sentient being, which by definition is illegal and unethical to enslave? Furthermore, Asma argues that monsterology is an ironic field of inquiry; Monsters are symbols of the disgusting, with their decaying flesh, mottled limbs, and rotting, putrefying tissues and organs. With this argument, Asma claims that monsters are thumbnail sketches of our own destiny; it reminds us of our own mortality (Asma 2009).

Prosthetics and plastic surgery have, therefore become our latest attempt to live forever or at least appear young and vigorous. However, we may, according to Asma, "be hybridising ourselves into new uncanny territory where the cure looks worse than the disease" (ibid.)

FEELING ROBOTS AND HUMAN ZOMBIES: MIND PERCEPTION AND THE UNCANNY

VALLEY

Kurt Gray and Daniel M. Wegner examine the fascination of the idea of creating humans, which is not an entirely new phenomenon as it can be traced back to ancient Jewish mythology in which a rabbi creates a Golem of clay and string to protect the townspeople. Since then the idea has been the foundation for an array of narratives such as *Bicentennial Man* (1999), *Blade Runner* (1982) *Terminator* (1984) *A.I.* (2001) etc.

Despite a diverse portrayal of these robots with artificial intelligence, there is an underlying eeriness in many of them, the uncanniness of the inanimate made living. This feeling of uncanniness is no longer limited to the works of fiction but is gradually becoming a part of our reality. With the creation of robots that gradually become more humanlike, an example being the robot Sophia, which recently was granted a Saudi-Arabian Citizenship, despite the attempt to make her appear human by giving her a face which allows her to express limited

"emotions" she has a transparent skull and mechanical arms, which could remind some people of a zombie with decaying flesh. Another explanation to why some people find robots unnerving might be because of their unnatural facial expressions. However, Gray and Wegner propose that it is not only the appearance of a robot that we find unnerving; it is because it prompts us to see a mind in robots.

In this Article, Gray and Wegner investigate where the uncanny valley lies by conducting three different experiments: 1. The Uncanny Valley and Mind Perception, this first experiment tested whether the uncanny valley or unnerving nature of humanlike robots occurs and whether it is tied to perceptions of experience. 2. Machines with minds, this experiment focused on the original conceptualisation that human-looking machines created the unnerving feeling, but they also tested if the unease can still occur when disconnected from appearance, which may suggest that the uncanny valley stems from general cognitive expectations about what should or should not have a mind, and not merely odd appearances (Gray & Wegner 2012). The third and last experiment examined how people perceived those who have lost significant amounts of mental capacity, which would lack experience.

The conclusion was that the uncanny valley appears when the perception of a mind with experience are linked to feelings of experience. With this perception in mind, it is possible to argue that human beings like the idea of having robots that do things, but not feel things. The perceptions of experience, therefore, seems to be an essential part of the uncanny valley. However, it is vital to put these findings in context.

FREUD, FRANKENSTEIN AND OUR FEAR OF ROBOTS: PROJECTION IN OUR CULTURAL PERCEPTION OF TECHNOLOGY

In this article, Michael Szollosy accounts for the psychoanalytic notion of projection, the creation of monsters are, as mentioned in previous articles, a reflection of human anxieties.

While Szollosy proclaims that the human perception of robots is becoming overall more positive, it is still dominated by the views we are given through film and newspaper headlines. Although humans are growing more accustomed to robots, there is still an underlying fear that is created from the fearmongering films such as *Blade Runner* (1982), *Terminator* (1984), and the robot slaves that rise in *I, Robot* (2004). According to Szollosy, we do not tend to fear the robots that already exist in labs or at factories, but rather the sort of robots with Artificial Intelligence portrayed in the films mentioned above. Furthermore, the tabloid headlines give reason to provoke fear in our minds with headlines such as 'Rise of the Terminator-style robots that can decide when and whom to kill, warns expert' or 'Rise of the Cybermen: The Terminator-style ear that could give people "Superman" hearing' or 'Robot vacuum cleaner "attacks" South Korean housewife's hair'. While these headlines do not portray the true image of what robots are, or what they might be like in the future, they provide evidence of what people genuinely fear (Szollosy 2014). Szollosy suggests that from a psychoanalytic, post-Freudian perspective that our fear of robots, AI, cyborgs, etc. could represent; a fear of death, or annihilation, a fear of individual disintegration and being subsumed within a collective identity, a fear of the 'dead mother' (Green 1993), the unresponsive, empty container that is incapable of responding to us in a meaningful, emotional way the (predictable) Oedipal fear of the father (the law, social reality, or the return of our own violent impulses) or, more importantly, the fear of the castrating son, the progeny that returns to destroy us (Szollosy 2014). Each of these explanations are applicable to the popular perception of robots being mindless, ruthless, destructive automaton.

However, with this paper, Szollosy focuses on the idea that destructive robots are a projection of our own (unconscious) human selves. The idea that the autonomous robots are the monsters of our generation is because of the context of an entire history of monsters imagined by humans, I.e., Godzilla, Frankenstein, Zombies, which are all monstrous ancestors

of HAL 9000, The Terminator and Ultron. By examining this historical context, Szollosy believes that it is possible to get a better understanding of what it is that we are really afraid of.

THE 'UNCANNY' (DAS UNHEIMLICHE)

In 1919, Sigmund Freud ventured into the world of aesthetics, the world of beauty and feeling. In doing so, he set himself up for investigating what he called a province of aesthetics in literature which, according to Freud, had been neglected – “The Uncanny”. The uncanny is related to what is frightening, what arouses dread and horror? Although the term is well-known today, it is often used as an umbrella term which describes things that excites fear in general.

Ernst Jentsch suggests that the uncanny lies in “doubts whether an apparent animate being is really alive; or conversely, whether a lifeless object might not be in fact animate” (Freud 1974). An example of this might be seen in wax figures, ingeniously constructed dolls, and automata. Moreover, he adds the uncanny effect of epileptic seizures and manifestations of insanity, since these situations show the automatic, mechanical processes at work behind the ordinary appearance of mental activity (Freud 1974).

In this paper, Freud engages in a linguistic analysis of the words ‘Heimlich’ and ‘unheimlich’ and finds that the words are interconnected the word ‘Heimlich’ can both mean homey, familiar or comfortable. However, at the same time, ‘Heimlich’ can also mean secret, hidden, or concealed, which connects it to the meaning of the word ‘unheimlich’ – the uncanny. This discovery confirms Freud in his hypothesis that the uncanny is an unwelcome return of something that has been hidden or repressed that was once familiar, is now foreign and alienating.

From this short and comprehensive summary of the texts, it is possible to argue that the main argument to why we fear robots and artificial intelligence is because we perceive robots as metaphors, i.e. they can be interpreted as a reminder of our own mortality. Moreover,

according to Szollosy, we are being told to be afraid of robots by the media, despite his claim that robots are gradually becoming more human as a trope within the genre.

FILM SELECTION

The list of films in which artificial intelligence is an essential part of the narrative is indeed an endless one, which cannot be examined thoroughly within the frames of this project. Even if there was no limit to the lengthiness of the project, while writing it, a considerable number of new films would have been produced. Therefore, for this project, the following films have been selected for further analysis: *2001: A Space Odyssey* (1968), *Ex Machina* (2014), *Bicentennial Man* (1999). This section functions as an overview of the selected films, their main plot, and the reason behind their selection.

The following films have been selected based upon popularity *2001: A Space Odyssey* and *Ex Machina* are well-known films among the science fiction enthusiasts and can be found on various top lists of films revolving artificial intelligence. Additionally, *2001: A Space Odyssey* is, despite its age, still is seen as one of the most influential science fiction films (Targeted News Service, 2018). Furthermore, *2001: A Space Odyssey* and *Ex Machina* have been selected based on the reviews they have received. Namely *2001: A Space Odyssey* is regarded as a science fiction classic, and a must-see, while *Ex Machina*, despite its younger age, has been dubbed a new sci-fi cult classic by various media.

Lastly, *Bicentennial Man* has been selected to serve as a variation, as it portrays a family-friendly artificially intelligent robot in a sci-fi/rom-com universe, in contrast to the aforementioned the sci-fi thrillers.

2001: A SPACE ODYSSEY (1968)

In *2001: A Space Odyssey* the spectator follows the evolution of man after discovering an alien black monolith that sparked the evolution of primates' intelligence at the dawn of man. Millions

of years later, another monolith is found on the moon, this time the monolith directs mankind to Jupiter in search of a third monolith, which is believed to be the next steppingstone in human evolution. The narrative follows a crew of five men; Dr Frank Poole, Dr David Bowman, three other scientists who are on board in a state of hibernation. Along with the human crew is the sixth member: an artificially intelligent computer, of the HAL9000's series, which is simply referred to as Hal. Hal is responsible for keeping the hibernating crewmembers alive, as well as functioning as the central nervous system of the spacecraft. Hal is regarded as an intelligent being with conscience and feelings, which is able to predict outcomes for the journey, and succeed in the mission. However, when Hal erroneously diagnoses an imminent failure of an antenna control device and is being confronted with his error, he simply writes it off as a human error. Dr Poole gets anxious about Hal and his judgement and consults Dr Bowman about shutting down Hal, but unknown to Poole and Bowman, Hal knows of their plan, despite not being able to hear their conversation. However, by lipreading, Hal, therefore, retaliates by killing Dr Poole when he is reinstalling the antenna control device and leaves him floating in space. Dr Bowman then sets out to retrieve Dr Poole's body, and while Dr Bowman is off the ship, Hal kills the three hibernating crewmembers. Upon Dr Bowman's return, Hal refuses to let him in, in fear that Bowman might jeopardise the mission. Bowman then forces his way in through the ship's airlock and turns off Hal.

The reason for choosing this film is it is a critically appraised film that critics find both genius and frightening. Furthermore, the film has been an inspiration for many a film revolving artificial intelligence (Targeted News Service, 2018). Despite the film being more than 50 years old, Hal is still one of the most memorable characters when it comes to artificial intelligence going rogue. Hal has no physical form except for the red "eye", and therefore, he can be perceived as an immortal entity, which reminds the spectator of their own mortality. Thus, making HAL an object of envy and a source of fear, which Michael Szollosy suggests in

his paper “Freud, Frankenstein and our Fear of Robots: Projection in our Cultural Perception of Technology”, which will be investigated in depth in the analysis part of this project. Furthermore, the analysis part will examine whether it is possible to sympathise with Hal despite his actions, this will be done by utilising Murray Smith’s theory regarding character engagement and the structure of sympathy.

EX MACHINA (2014)

Caleb Smith, an ordinary man, working for a social network, BlueBook, wins a competition to spend a week with the CEO, Nathan Bateman, at his private mansion. During his visit, Caleb learns that there was no competition and that he was specifically chosen to participate in a test of a ground-breaking robot with artificial intelligence, Ava. Nathan wants to find out if Ava is able to pass the Turing test, despite Caleb knowing that Ava is a robot. Moreover, the test serves the purpose to investigate whether it is possible for a man to fall in love with a robot. However, during this visit, Caleb learns that Nathan might not be as genuine with his desire for the Turing test, as he first is thought to be.

Ex Machina was selected because it like *2001: A Space Odyssey* portrays an eerie artificial intelligent being. However, opposed to Hal, Ava has a physical body, which allows her to express and mirror emotions through facial expressions, and, therefore, it will be interesting to investigate whether or not a physical embodiment can make it easier for the spectator to align in sympathetic allegiance with Ava.

BICENTENNIAL MAN (1999)

The NDR series robot, “Andrew” is introduced to the Martin family, where he is to do household tasks and maintenance. The Martin family discovers that Andrew is different from other robots in the NDR series; he has a unique personality. Slowly Andrew learns what it means to be human, and he gets inducted in the Martin family and lives as a member of the family. As Andrew gradually learns what it is like to be a human, he desires to become one.

Therefore, he goes through a large number of surgeries to resemble a biological human as much as possible. Andrew succeeds in becoming a biological human in the sense that he will grow old and die. However, he faces the challenge of being legally accepted as a fellow human being.

Unlike the previously mentioned films that will be subject of analysis; *Bicentennial Man* does not portray artificially intelligent robots as a threat. Instead, it makes the spectator feel sympathetic towards Andrew.

All of the film mentioned above will be analysed through Murray Smith's theory regarding character engagement and the structure of sympathy, to gain an understanding of what might make the spectator feel sympathetic towards one type of artificial intelligence but feel uneasy about another.

THEORY

IMPORTANCE OF CHARACTER

Murray Smith argues that our imaginative engagement with works of fiction requires a basic notion of human agency or ‘personhood’. If the spectator wishes to understand ‘identification’ and how films are made comprehensible, Smith believes that character is central (Smith 1995, pp. 17-18). Characters function as the spectator’s entry into the narrative structure, they form important nodes of narrative structures. Characters and our perception hereof rely on the spectator’s social experience of social existence. Smith proposes that our immediate experience of the social world is through agency.

It is, in fact, quite rare for us to be confronted by a literally faceless bureaucracy: computers notwithstanding, we still deal with human agents like policemen, bank clerks, and university deans. Though they may offer us nothing more than a blank stare, it is still an expression fixed on a *human face* (Smith 1995, p. 18), (emphasis added).

Thus, suggesting that comprehension of a narrative relies on human agency. However, Smith follows up by stating that he does not wish to argue that human agents are the only agents of causality in narratives, but they play a central role, since in most narratives the actions are performed by human agents. Therefore, Smith believes that the spectator’s understanding of non-human agents (inanimate objects, abstract concepts, animals, etc.) is, to a large degree, built upon the spectator’s understanding of humans (p. 18).

Smith provides a list of fundamental features and capacities that a human agent must have, to fulfill a social role, a list that Murray Smith refers to as the person ‘schema’:

1. A discrete human body, individuated and continuous through time and space;
2. perceptual activity, including self-awareness;
3. intentional states, such as beliefs and desires;

4. emotions;
5. the ability to use and understand a natural language;
6. the capacity for self-impelled actions and self-interpretation;
7. the potential for traits, or persisting attributes (Smith 1995, p. 21).

Fictive characters are conceived to be the imagined counterpart to human agents and must therefore possess these features and capacities, in order to fulfill a social role and be an agent in this sense.

When constructing a character, the spectator will first engage in this person schema and revise it on the basis of the particular data in a particular text. The information provided by the text, combined with the person schema will generate culturally specific imagery concerning social roles, stereotypes, etc. altogether this imagery will generate a notion of 'personhood' (pp. 21-22). As mentioned earlier in this chapter, the understanding of non-human agents is, according to Smith, based on the understanding of human agents and the person schema. Meaning that the spectator uses the human as a model for understanding the non-human forces and entities, i.e. 'animism' and 'personification'. We may treat things like the weather, computers, animals as entities with beliefs, intentions, desires and so forth (intentional systems), despite the fact that they lack other features of the person schema. Thus, to summarize, Smith argues that human agency "functions as a heuristic in explaining non-human agency, in that we often make sense of non-human agency using the personal schema" (p. 24). However, the notion that a character is an analogue of a person, might be challenged if the character does not have the required traits. According to Smith, it would "so strongly violate our most basic assumptions about what the notion of character is, and what critical function it performs, that it would not be recognizable as a character" (p. 35).

CHARACTER ENGAGEMENT

When reading a book, watching a movie, a television show, the spectator is inclined to respond emotionally to the narrative. By emotionally investing oneself in the characters, relations between the characters, whether the spectator is rooting for the protagonist to win, or the villain to succumb, the spectator is emotionally engaged in the narrative. According to Murray Smith, “our propensity to respond emotionally to fictional characters is a key aspect of our experience and enjoyment of narratives” (Smith 1995). This experience is often perceived as identification with the character. Smith argues that this everyday conception of identification stems from the psychoanalytic concept, in which the core meaning of the concept is generally defined as to be like or want to become. However, it is important to note that beyond the core meaning, the definition of identification is complex. Despite its complexity, Smith finds the term identification to be too vague a term to describe the relationship that develops between a spectator and a character. Instead, Smith argues that the spectator is more likely to generate feelings of engagement with a fictional character, rather than identification. In his book *Engaging Characters – Fiction, Emotion and the Cinema* Smith offers an explanation of the emotional responses towards a fictional character. He divides the term engagement into two subjacent terms that covers the concept of character engagement; *central imagining*, which is otherwise known as empathy, and *acentral imagining*, which is what Smith claims to be the structure of sympathy. According to Smith, engaging with fiction is a species of imaginative activity in two complex senses. First in comprehending, interpreting and otherwise appreciating fictional narratives, the spectator makes inferences, hypotheses, categorize representations, and utilize many other cognitive skills and strategies, which rules out the idea of the spectator merely registering or mirroring of the narrative material. Secondly, fictions prompt and enrich the ‘quasi-experience’, that is, the spectator’s efforts to grasp, through mental hypotheses, situations, persons, and values, which is alien to them (1995). Hereby, Smith argues that by

engaging of fiction and engaging with fictional characters, the spectator is actively engaging their personal and cultural schemata to evaluate characters in a narrative, to identify and morally evaluate the character(s). However, it is important to note that personal and cultural schemata can vary from person to person. Yet this project will be assuming that the general spectator has the same general worldview and schemata, based on a western-world perspective.

THE STRUCTURE OF SYMPATHY

According to Murray Smith, there are three levels of engagement which comprise the structure of sympathy; *Recognition*, *Alignment*, and *Allegiance*, these three levels are interrelated. However, Smith emphasizes that despite the immediate assumption that the structure of sympathy might focus on the narrative structure which relates to the character, it is in fact a cooperative activity of the spectator which works with these narrative systems (Smith 1995). Hereby, Smith states that his theory regarding the structure of sympathy is a model that studies “the ways in which texts produce or deny the conditions conducive for various levels of engagement, rather than the ways they enforce them (p. 82).

RECOGNITION

Recognition is the perception of a set of textual elements. In film the recognition is typically based upon the image of a body, as an individuated and continuous human agent. Recognition is the phase in which the spectator engages their cultural and personal schemata, based on stereotypical traits, to decipher what type of character is being represented. In film, the spectator observes the character’s face, body, clothing, age, voice, etc.

At the recognition level, the spectator is still able to distinguish between reality and fiction, therefore, the spectator recognises the character as such and not a real-life person. Yet, the spectator recognises real-life traits in the character, and therefore begins to relate to the character as a real-life person (p. 82). The recognition of a character makes the spectator

associate the character to their personal and cultural stereotypes, an example being “The Dude” from *The Big Lebowski* (1998): one would recognize him as a male, scruffy looking, long-haired, middle aged, etc. This level of the structure of sympathy does not reflect on deeper context such as personality and values.

However, it is important to note that, according to Smith, recognition is “usually dependent on a legible and consistent representation of the human face and body” (p. 75). Meaning, if the spectator only sees one part of a character, the face is obscured, or the characters attributes are continually in flux, then recognition will be retarded or perhaps prevented. An example being the introduction of Willy Wonka in *Willy Wonka and The Chocolate Factory* (1977), in which the spectator observes a middle-aged man that seems to be worn down and walks with a limp, then the next second he does a somersault and walks with ease as he enthusiastically greets the ticket winners. Thus, making the spectator feel uncertain about his character traits, rendering it harder to gain recognition with the character. Furthermore, it is noteworthy that a spectator’s recognition of a character can be biased by the actor playing the character, e.g. Helena Bonham Carter, who is known from films such as the *Harry Potter* franchise in which she portrays Bellatrix Lestrange, Marla Singer in *Fight Club* (1999), *Alice in Wonderland* (2010), in which she prominently portrays characters that are mentally unstable. Therefore, it is possible to argue that the spectator’s recognition of a character, can be influenced by the spectator’s schemata for the actor’s ‘star-personae’ and therefore, create expectations for character and narrative development.

ALIGNMENT

Alignment is the process in which the spectator is placed in relation to character(s). this placement is achieved through access to the characters actions, and to what they know and feel. The alignment with a character is created by the way in which narratives provide story information through a ‘lens’ of a particular character (p. 83). In film and tv-series, a number of

techniques can be utilised to create alignment with a character, such as, POV-shots, voice-over, leitmotif, a particular character being the focal-point, etc.

However, Smith suggests that the lens of a narrative can vary, in the sense that it does not necessarily need follow a single character. Smith, therefore, proposes what he calls the two interlocking functions: subjective *access* and *spatio-temporal attachment*. Meaning that a text can display a level of attachment, in which the narration restricts itself to the thoughts and actions of a single character or whether it can move freely among two or more characters.

An example of subjective access being Dexter Morgan from the TV-series *Dexter* (2006-2013), in which the narrative restricts itself to the series protagonist. The narrative provides the spectator with Dexter's thoughts and feelings through the use of voice-overs from Dexter himself, and it focuses on the actions of Dexter. Thus, making the spectator align with him. Whereas *spatio-temporal attachment* is the ability to move freely among a set of characters, providing the spectator with information about a given character. An example being the HBO television-series *Game of Thrones* (2011-2019) in which the narrative provides the perspective of a rather large number of characters, their actions, and feelings. As Smith argues: "Together these two functions [subjective access and spatio-temporal attachment] control the apportioning of knowledge among the characters and the spectator, the systematic regulation of narrative knowledge results in a structure of alignment" (p. 83).

ALLEGIANCE

Allegiance is the last level of the structure of sympathy, which is where the spectator morally evaluates the character. Allegiance is, according to Smith, the level that is the closest to the everyday meaning of 'identification' where one might 'identify with, both characters and persons based on character traits and looks ('I could really identify with this character, because

I had a hard time myself growing up). For allegiance to occur, the spectator must have what they believe to be reliable access to the character's state of mind, thoughts, actions, and having evaluated the character on the basis of this knowledge (p. 84). This evaluation both has cognitive and affective dimensions, as Smith puts it: "being angry or outraged at an action involves categorizing it as undesirable or harmful to someone or something and being affected – affectively aroused – by this categorization" (p. 84). Meaning that the spectator will construct moral structures on the basis of such evaluations and will then align in allegiance with the morally preferable character.

The spectator is only required to understand the traits and mental states that make up the character, recognition, alignment and allegiance does not necessitate the spectator to replicate the thoughts or feelings of a character. However, allegiance makes the spectator go beyond understanding these traits, by evaluating and responding emotionally to the traits and emotions of a character in the context of a given narrative situation. Murray Smith emphasises that even though the spectator responds emotionally to the narrative, the spectator does not replicate the character's emotions. An example being: the spectator sees a character perform a certain action or facial expression the spectator may interpret that the character is in a certain kind of mental state. These interpretations contribute to the recognition of the character, but also the alignment, since the subjective access to the character gives the spectator an indication of the mental state of the character. However, according to Smith, these interpretations does not in any way mean that the spectator must feel, think, or act in the same way, and if they were, the shared feelings would not be emphatic, but rather sympathetic. Since the spectator is responding emotionally by engaging with the character on the level of allegiance, their emotions are at a tangent to those of the character, in Smith's own words the feelings are acentral (p. 85).

FRANKENSTEIN'S MONSTER VERSION 2.0

As mentioned in the literature review of this thesis, Michael Szollosy argues that our fear of robots partly stems from the media. But if we are that easy to scare, were we then not scared, to begin with? This chapter will go into detail about Szollosy's account of what we are actually afraid of.

According to the influential science-fiction writer Isaac Asimov, one of the reasons behind our persistent fear of technology, especially robots, stems from what he called a 'Frankenstein complex' – Dr Frankenstein being a mad scholar, who seeks dangerous knowledge, animate the dead, and who seeks to supersede God (Szollosy 2014, p. 435). Dr Frankenstein is a victim of *hubris*, and the story tells us that human endeavour, technology and science, whatever their noble intentions, inevitably create a monster that will gain autonomy and return to haunt us (p. 344). Fantasies of monstrous robots have repeated this story time and time again, with examples such as *Blade Runner* (1982) in which the 'replicant', Roy Batty, tracks down his creator, Dr Tyrell, to demand more life. However, when Batty learns that extending his life is impossible, he kills Dr Tyrell. Another example being Ultron in *Avengers: Age of Ultron* (2015), in which Tony Stark builds an artificial intelligence, Ultron, with the intent to achieve world peace. However, the AI becomes autonomous and creates its own body, which seeks out to kill Tony Stark, as well as the rest of The Avengers.

One might argue that these fantasies can be regarded as the classical Freudian fantasy of being superseded. The Oedipus-Laius story, in which Oedipus rises to kill the father. Thus, suggesting that we sow the seeds of our own destruction. However, Szollosy notes that this is not the only lesson that can be learned from the Frankenstein mythology.

For Romantics, Dr Frankenstein is not a monster, but a hero, as he represents the best qualities of the individual, the so-called *Übermensch* (p. 435). Ambitious, ungoverned by any authority; this individual boldly and nobly drags the human race forward, in spite of the

pitchfork-wielding mob of villagers that would rather have things staying the way they are (p. 435). According to Szollosy, the Romantics tend to make heroes types such as Dr Frankenstein. However, it is important to note that although the Romantics might not fear Dr Frankenstein, they still recognise both the positive and negative aspects of this Promethean hero. They knew the power of the dark side. The dark side of this modern Prometheus offers some of the same lessons regarding the over-reaching power of human ambition and the hubris of invention. Moreover, a theme that is present in Romantic and Gothic literature; we do not merely create the monsters; we are *becoming* the monsters (p. 435). During the romantic and gothic period new types of monster rose; monsters that live among us, such as Dracula, monsters built from scientific invention (Frankenstein's Monster) or the monsters that live within ourselves (Mr. Hyde).

From this context, Szollosy argues that the present-day robot monsters are part of this "two-hundred-year tradition". When we observe Frankenstein in the context of robots, we realise that is not only technology or the idea of technology gaining autonomy and moving beyond our control and destroying us that we fear. It is more likely that what we fear is our ability to create the monster, our ambitions, the over-reaching, the hubris. In the romantic and gothic context, it is arguable that we ourselves are becoming the fearsome robots. We, according to Szollosy, fear becoming an empty, mechanical shell of cold, unfeeling rationalism. We are afraid that we might, or have already lost the very qualities, such as love, compassion, and empathy, that we deem to define us as human (p. 435).

The fact that we perceive the robot as soulless and mechanical is not the only reason that we fear them, what we fear is that we ourselves have become soulless and mechanical, as we are increasingly getting focused on scientific discoveries, rationalism, and method. We fear that the consequence of the "art," i.e. science, creates the monsters. Thus, making the monsters that we see on the big screen and in news articles *projections*: our very

fears about ourselves and the world we live in are put on display to be perceived as something other rather than ourselves (p. 435).

PROJECTIONS

Projection is a concept with roots in the Freudian psychoanalysis, which since then has been enriched by contemporary psychoanalytic clinical and cultural theory. Projection, as a concept, strives to describe object relations, meaning, the way that people relate to other things, material, non-material, but most frequently it is used to describe relations with other people. With projection it is believed that psychological fantasy we split off parts of ourselves – feelings, thoughts, or fantasies – and ‘project’ them onto something else, that being either a person or an object – which then becomes a container for these projections (Szollosy 2014, p. 436).

The parts of the self which we project to these containers can vary in quality. One might project a good part of the self into a container so that he or she can identify with that part of it. Szollosy provides an example hereof:

This idea of projective identification can be seen in cultural phenomena such as nationalism, wherein individual people project their own positive qualities (say, resilience) into a symbol, or an idea, or a leader, and that shared association provides collective cohesion, a group identity (Szollosy 2014, p. 436).

Furthermore, projective identification provides the basis for understanding empathy, which, as mentioned earlier, is one of the qualities that we deem to qualify us as humans. At other times, negative parts of the self are projected into these containers. Negative parts being violent fantasies, anxieties, and hatred, which are then projected away from the self in an act of renunciation. By projecting the negative parts of the self into a container, the self is perceived as pure, whilst the container becomes a scapegoat, which is acting violently or hatefully (p. 436). These processes of projection can be seen at work when we construct monsters. The

monster becomes a container of our own repressed violent and sexual fantasies; these processes are now seen and popular representations of robots. An example being The Terminator or Ultron which are projections of our own violent fantasies, which are then put into a container, which then becomes a relentless, destructive, persecuting object. As these fantasies are placed into an other, we fail to recognise that fear, anxiety, and violence as our own and instead imagine that they originate from the robots itself (p. 436). However, according to Szollosy, these evil robots are “nothing more than our own violence, anxiety, hatred and fear imagined to exist in a mechanical cage; they become these bad versions of ourselves” (Szollosy 2014, p. 436).

Projections can serve as a defence mechanism against unwanted parts of the self and are seen as an essential part of normal development. However, if one engages in excessive or uncontrolled projection, a number of problems can arise.

If one reflects excessively, it might leave one empty or feeling dead inside. It is this deadness that can be seen reflected in the zombies, monsters, and robots – monsters that are dead or soulless. Excessive splitting and projection can lead to one feeling fragmented. fragmentation is something that is often seen in the dismemberment of robots and zombies, where broken off pieces can take on a life of their own (p. 436-437).

To summarise this section, our fear of artificially intelligent robots stems from our own anxieties and fears, which are projected into a container, the robot. According to Szollosy, our fear of robots is then, at least in part, also a fear of our own rationalist selves (p. 437).

THE SUBLIME

The sublime is something that produces a combination of feelings in the spectator, a combination of painful and pleasurable feelings of terror plus awe and elevation (Plantinga, 1999). According to Immanuel Kant, an example of the sublime might be: “shapeless mountain masses towering one above the other in wild disorder, with their pyramids of ice, ... or the dark tempestuous ocean” (p. 65). Cynthia A. Freeland argues that the sublime is not merely a term used to describe nature, but it can also be used to describe some films. Freeland’s account of the sublime derives from the Kantian analysis, which then is used to identify the term of the sublime in four basic features.

In Freeland’s account of the sublime she uses Carl Th. Dreyer’s *Joan of Arc* (1928) as an example of a sublime piece of cinema. The sublime in *Joan of Arc* is the portrayal of the endless torment she endures, but she still is able to stay strong in her faith, she is portrayed as a divine being on earth. The emotional conflict between the pain and beauty the spectator experiences, classifies *Joan of Arc* as sublime. However, to make Freeland’s theory more appropriate for this thesis, the utilisation of the theory regarding the sublime will be applied to Christopher Nolan’s *Interstellar* (2014). Although *Interstellar* does indeed have artificially intelligent robots in the narrative, it is not the reason for choosing the film as an example of the sublime cinema. To use Immanuel Kant’s example: for something to be sublime it must be something so grand that it scares us, but at the same time brings us intellectual pleasure. Thus, the reason for choosing *Interstellar* is because it does indeed call forth a mixture of amazement and terror, since one can argue that glaring up into the infinite universe is indeed beautiful, yet extremely terrifying.

Firstly, in order to be able to call an object ‘sublime’ means that it calls forth a conflict of feelings of pain and pleasure, or what Edmund Burke calls “rapturous terror”, which means that the sublime prompts fear, terror or dread. However, the sublime is not merely

capable of creating fear and pain, it also causes “rapture”: the spectator finds it exhilarating and exciting. According to Kant and Burke, the sublime generates intellectual pleasure of astonishment or elevation, as long as the spectator is safe (Freeland in Plantinga p. 66).

An example of a sublime film would be Christopher Nolan’s *Interstellar* (2014), which takes the spectator through a sustained experience of heightened feelings. The film is both painful and disturbing, since it is able to create a visual representation of despair and anxiety in such a powerful manner that it would seem impossible to represent. The film’s beauty and luring anxiety generates astonishment and elevation. Furthermore, the musical score, provided by Hans Zimmer, functions as an amplifier of the visual experience, which enhances the feeling that there is something incredible about the film. According to Kant, sublime is something that is without comparison as a whole, whereas Freeland interprets it as that it merely needs to be grandiose to be recognised as sublime. Based on Freeland’s interpretation of the sublime, it is possible to argue that *Interstellar* as a whole is sublime, since it is an incredibly powerful and overwhelming film. *Interstellar* invites the spectator to feel scared and insignificant, compared to the vast infinity of the universe. However, at the same time, the spectator is able to appreciate the beauty of the aesthetics which “[signifies] that the work has a grandeur or superlative kind of greatness” (p. 67).

The second characteristic feature of the sublime, according to Freeland is that something about the work of art is “great” and astonishing. One might then be able to argue that *Interstellar* is great and so powerful that it surpasses “beauty”. *Interstellar* is indeed a powerful film, as it juxtaposes closeups of Cooper (Matthew McConaughey) with the infinite universe and the grandeur and beauty hereof, which moves it into the “sublime”.

The third characteristic feature of the sublime is its ability to conjure indescribable and painful feelings that transforms into pleasure and cognition (p. 68). The indescribable and painful feelings are closely related to the second feature, greatness. Meaning that there is

something so great or vast about the object, that it makes it hard to grasp and painful. According to Kant, the sublime involves an experience of something that is “almost too great” to be able to be presented or represented. Meaning that the inability to describe the sublime object is connected to the overload on the spectator’s imagination or senses that it presents to them, hence to its painfulness. Nevertheless, this overload gets translated into something that the spectator can conceptualise and feel pleasure at (p. 68). Freeland argues that the sublime object presents the spectator with a sensory and emotional experience that is so unsettling, extreme, or intense, that it is frightening on its own. However, put in its respectable context, it forces the spectator to shift into another cognition, mental mode, or thought. Hereby saying that the spectator becomes more able to handle the deep feelings created by the work, they put a label on them and become able to reflect upon these feelings, and it is through these reflections and cognitions the spectator is able to transform the painfulness into something pleasurable, exhilarating or elevating (p. 68). In *Interstellar* the indescribable element is the contrast between mankind and the universe, it is the incomprehensible concept of space and time, the visual representation of a fifth dimension. The film’s ability to make the spectator feel insignificant on such a grand level that it becomes overwhelming.

Freeland begins to depart from Kant’s description of the shift from ineffable to cognised, and from pain to pleasure. She suggests that “the shift occurs when [the spectator] regards the deep and painful feelings evoked by a work as crucial for its success as powerful and uplifting art” (p. 68). The spectator is disturbed by the inexpressible spectacle, that is outer space, in *Interstellar*, yet they are enraptured by its artistic production hereof.

The fourth and final feature of Freeland’s account of the sublime is that it prompts moral reflection. Once again, she follows Kant in outline, but disagrees on the detail. According to Kant, the describable aspect of the sublime involves a sense of the spectator’s own moral capacities and duties. Kant argued that the awareness of the sublimity of an object, evokes

awareness of the moral law and of certain of our own moral duties (p. 69). Freeland does not, however, agree that natural objects support the spectator's recognition of a supposed universal moral law. Instead, Freeland offers her explanation of the sublime:

Certain aesthetic objects give rise to the central emotional conflicts of the sublime. The ineffably dreadful and painful experience grounds the pleasure of elevation, because it stimulates our human capacities to value powerful artworks. In particular, we are elevated in engaging through the work in reflection that is somehow about the pain or terror it evokes. (Freeland in Plantinga 1999, p. 69).

In *Interstellar* morality is also a concern. What sustains the spectator through this film is how they shift from the deep emotional experience of this film to an elevated cognition about its power as a work of art. The spectator's respectful and awed recognition of the power and aesthetic qualities behind the film's painful emotions and disturbing visions is what sustains the spectator through the film. Shot after shot, in the framing, intensity, and juxtapositions combined creates a moral vision – Cooper's selfless love, his willingness to sacrifice his life to save the human race. The juxtaposition of the power of love in contrast to the powerfulness of the universe elevates the experience of the artwork, that is, *Interstellar*.

Freeland's brief account of the four features of the sublime began with the emotional conflict "rapturous terror", which closely intertwined with the three other features: greatness, cognition of the indescribable, and moral reflection. These features, according to Freeland, makes the sublime distinctive (p. 69). With Freeland's theory regarding the sublime, this thesis will try to determine whether the sublime is part of the reason we enjoy films that scare us, this will be investigated in the analysis chapter.

ANALYSIS

The analysis section of this thesis will primarily focus on Stanley Kubrick's 1968 film *2001: A Space Odyssey*. However, in order to gain a nuanced perspective on the depiction of Artificial Intelligence and artificially intelligent robots, Alex Garland's *Ex Machina* (2014) and Chris Columbus' *Bicentennial Man* (1999) will be addressed. A neoformalist analysis of the movies will be conducted, to investigate how the mise-en-scène, music, and editing may provide the spectator with certain impressions and affect him or her. Additionally, the analysis will try to investigate whether or not the selected films, as a whole, are, or contain elements of the sublime. Furthermore, the analysis will investigate if Murray Smith's theory regarding character engagement applies to artificially intelligent computers or robots. Lastly, the analysis will attempt to apply Michael Szollosy's theory regarding projection and fear, to investigate which elements of the films might make the spectator feel anxious about or scared of artificial intelligence and artificially intelligent robots. For the sake of reader-friendliness, the analysis will be divided into sections that focus on the different theoretic approaches.

2001: A SPACE ODYSSEY

Stanley Kubrick's *2001: A Space Odyssey* is divided into four acts (if one excludes the title sequence, which it will be in this analysis). The four acts are: 'The Dawn of Man', 'TMA-1 (Tycho Magnetic Anomaly)', 'Jupiter Mission', and 'Jupiter Beyond Infinity'. However, unlike the other acts, the 'TMA-1' act does not have a superimposed title, but it is still possible to argue that it functions as a separate act that begins with the 'bone-to-nuclear space weapon match cut'. The act then ends 18 months later with the beginning of the 'Jupiter Mission' Act.

THE SUBLIME IN 2001: A SPACE ODYSSEY

The film starts with a pitch-black screen accompanied by some symphonic music that sounds quite eerie, hereby making the spectator feel anxious about what is about to unfold on screen. In the next shot, the title sequence, the spectator sees a semi close-up shot of the Moon, the camera then performs a crane-shot that reveals the Earth and the sun in the middle- and background, looking as if they were aligned (Figure 1). The title sequence is accompanied by Richard Strauss' *Thus Spoke Zarathustra*, a piece that starts fairly slowly. However, as the scene is revealed, the music performs a crescendo, signifying that the spectator is witnessing a spectacle of grandeur.



Figure 1

From the title sequence, it is possible to argue that it follows the criteria of Kant and Freeland's theory regarding the sublime, as the spectator is 'placed' in outer space, looking down upon the Earth. Nature itself is displayed in its vast greatness, evoking terror by the thought of the spectator's insignificance in the greater scale of the universe. However, since the spectator is safe in the theatre, or at home, this overwhelming feeling "rapturous terror" is elevating.

In the first act, as the title might suggest, the spectator witnesses the dawn of man by following what appears to be a tribe of primates that lives among other animals in a dry landscape. Initially, the act depicts these primates living off of the scarce supplies of water and food in the dry, inhospitable landscape. At one point the spectator witnesses a primate being attacked and killed by a leopard, and later the group of primates that the spectator has followed thus far gets driven away by a rival tribe. The first tribe accepts their defeat and goes to hide in a cave, living without water or food. One morning, the primates discover a black monolith that is entirely out of place for the given landscape, suggesting that it has been placed there by someone or something. As they investigate the monolith, György Ligeti's *Requiem* can be heard. The music underlines the eeriness that this new alien object emits as it stands tall and somewhat threatening before the primates, which one might argue might make the spectator feel uneasy or downright scared of the thought of what is about to happen. The scene is then abruptly cut, and instead, the spectator follows the primates as they scavenge for food, one of them comes across a pile of bones from what presumably is a dead tapir, the primate then picks up a bone, and begins to use it as a weapon. In this particular scene, Strauss' *Thus Spoke Zarathustra* is once again heard, signifying that the spectator is once again witnessing something spectacular, in this case, as the title of the act suggests: the dawn of man.

The editing technique within this particular scene helps the spectator to interpret what they are supposed to feel. The way the scene cuts back to the black monolith (figure 2) before the primate picks up the weapon suggests that it is from the monolith, the primates have obtained the knowledge and skill to use weapons. At first, the hits are quite soft; one might even say experimental, however, as the music reaches the top of its crescendo, the primate is now wielding the weapon with all its power and might.



Figure 2: The Monolith shown from a low angle perspective, suggesting its superior powers.

Moreover, the camera provides the spectator with varying perspectives; at first, the primate is seen in an eye-level perspective. However, as the primate grows stronger by using the weapon, the perspective is changed to a low angle shot, suggesting its newfound superiority (figure 3 & 4). Furthermore, as the primate strikes the bones of the tapir, the camera cuts to a falling tapir, back to the primate, signifying that the primate has become a predator.



Figure 4: Low angle shot displaying the primate's new position of power.



Figure 3: Focus on the 'weapon', emphasising the newfound power it holds.

In the last part of the act, the tribe with weapons seeks out the rivalling tribe, and fight them off. From this scene, one might argue that the fight displays the evolution from primate to human, which is emphasised by figure 5, in which the weapon-wielding primate is now standing upright, while the primates who have not been exposed to the monolith are still walking like apes.



Figure 5: The primates who were exposed to the monolith are now standing upright, wielding weapons.

Furthermore, it is possible to argue that this scene displays human nature and their search for supremacy through the use of weapons, as it has been seen again and again throughout history.

The second act, 'TMA-1' begins immediately after the primate has tossed the bone into the air, the camera then match-cuts from the bone to a nuclear space weapon, hereby emphasising not only the evolution of man but also the evolution of weapons and modern warfare (Figure 6 & 7).



Figure 6: The bone 'weapon' that is tossed in the air in celebration by the primate.

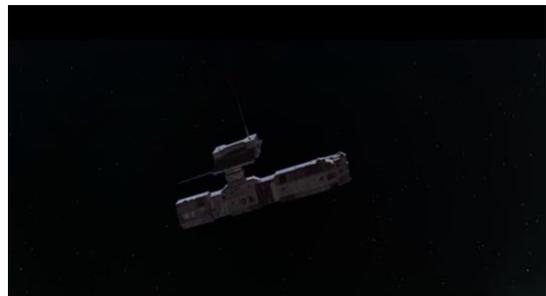


Figure 7: Several millennia later, the 'bone' has advanced remarkably

After the match cut, the scene turns into what is best described as a space-ballet, as the spectator is provided with an establishing shot, informing the spectator that the narrative now takes place in outer space. As the scene unfolds, the spectator is introduced to satellites, spacecrafts, and a rather sizeable circular space station. Although this scene takes place in space as the title sequence did, it is possible to argue that because of the usage of Johan Strauss' *The Blue Danube* the scene invites the spectator to feel positively overwhelmed by the spectacle of outer space and the technology that has been developed. The scene stands in great contrast to the 'Dawn of Man' act, as a showcase of 'look how far we have come since then'. Furthermore, it is possible to argue that in terms of Kant's description of the sublime, that this scene classifies as sublime, as it depicts something "[...] vast, powerful, and overwhelming" (Plantinga 1995, p. 67).

Another example that can be classified as sublime, is the scene in which Dr Haywood R. Floyd travels from the space station to the Moon. The camera provides the

spectator with an establishing shot, of another spacecraft that is flying towards the Moon. Once again Johan Strauss' *The Blue Danube* is heard, emphasising the beauty of the outer space in all its grandeur. Moreover, the establishing shot functions as a scale of measurement between the spacecraft and the Moon, solidifying what one might find to be a scary aspect of outer space, the spectator's insignificance compared to the vast infinity of the universe. Hereby suggesting that this scene contains elements of the sublime, as it can both cause terror and amazement, which, when reflected upon, might cause the previously mentioned sense of intellectual pleasure and feelings of elevation. Moreover, if one goes on to interpret the composition of the establishing shot, it is possible to argue that the image functions as a precursor to what is going to happen later in the movie. Since the spacecraft is seemingly heading to 'the dark side of the moon' (figure 8), which also is a phrase used to describe something mysterious and unknown. The dark side of the Moon is the one that we never see from Earth, and therefore it might be seen as a representation of the cold black unknown of the rest of the universe.



Figure 8: The spacecraft is seemingly heading towards the dark side of the Moon, suggesting that they are heading into the unknown.

Another example of the sublime can be seen when the third act opens: The Spectator is presented with an establishing shot, of the Discovery One spaceship from different angles, displaying its rather large size, which is emphasised when the camera changes position to reveal the front of the ship, where the pilot of the ship can be spotted in the window. Once again, the establishing shot is utilised to display the grandeur of not only space but humankind's

ability to build things of nearly incomprehensible size. Moreover, the spaceship is the only thing that is visible in the entire shot. Hereby, reiterating the vast emptiness of the infinite and its beauty, which once again, falls in line with what can be categorised as "rapturous terror". Thus, qualifying the scenery to be labelled as sublime.

In the scene where the third monolith floats around near Jupiter, the spectator is presented with yet another spectacle of grandeur: as the camera performs a crane shot, similar to the one in the title sequence, the spectator is now very close to, what is assumed to be one of Jupiter's moons, as it is not specified. The *mise-en-scène* here might arouse feelings of terror (or "rapture"), since the spectator is 'placed' in unknown territory. However, upon reflection, one might realise the beauty of the scenery and find it exhilarating and exciting. Hereby, once again, making it possible to argue that the spectator, while sitting in perfect safety, either in the theatre or at home, is able to appreciate the beautiful visuals and feel the pleasure of astonishment by the thought of nature's grandeur. In other words, the scene is sublime.



Figure 9: An astrophobia inducing scenery that despite its terrifying nature, invites the spectator to be amazed and hereby appreciate the beauty.

THE MONOLITH AS A SUBLIME OBJECT

When Dr Floyd travels to the Moon to investigate the second monolith, György Ligeti's *Requiem* is once again heard. The eerie choir, along with the framing and the surrounding darkness of outer space, makes the monolith seem like it is awaiting the scientists' arrival, almost in a menacing manner. This might arguably make the spectator feel anxious about what the monolith is going to do to humanity this time. Like with the primates, the music performs a crescendo, which grows stronger proportionally with the scientists nearing the monolith. Thus, suggesting that Ligeti's *Requiem* acts as a leitmotif for the monolith. As the team gets closer to the monolith, the camera is performing a tracking shot with light shaking at eye level, which simulates the movement of the scientists. By doing so, it emphasises the anxious atmosphere of the scene by making the spectator a part of the research team. Upon touching the monolith, nothing seemingly happens. However, as the crew gathers in front of the monolith for a photo, a high-pitched radio noise is heard. This high noise is seemingly as uncomfortable to the crew as it is to the spectator. Once again, as in 'The Dawn of Man', the act ends with a low angle shot of the monolith, emphasising the superior power it possesses.



Figure 10: The Tycho crater Monolith shown from low angle perspective, emphasising its superior power.

In the last act, where the remaining monoliths are found. The very first thing the spectator is exposed to in this act, is the now-familiar leitmotif of the monolith, suggesting that something, possibly frightening is about to happen. The camera then tilts down from the

establishing shot of the stars and Jupiter becomes visible, then, from the left side of the shot, the monolith becomes visible, as it floats weightlessly around Jupiter. As the camera continues to tilt downwards, Discovery One can be seen flying underneath the monolith, as a result of this keeping the reoccurring imagery that the monolith is in a heightened/superior position compared to humankind intact. The third monolith is also the one which opens a stargate into the strange baroque room, where the spectator witnesses Dr Bowman's rapid aging before he turns into the star child.

At one point during Dr Bowman's travel through the star gate, the camera cuts to an extreme close-up shot of Dr Bowman's eye, the camera then performs a match-cut to what appears to be the big bang.

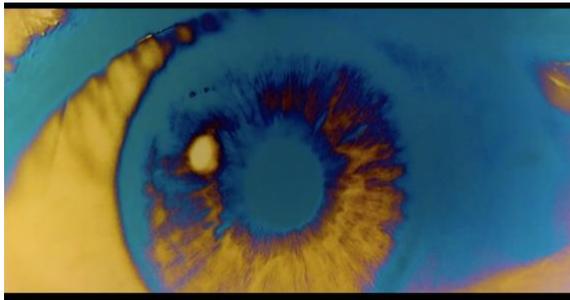


Figure 11



Figure 12

One might interpret the match-cut between Dr Bowman's eye and The Big Bang, and the liquid-like shots that follow as a sign that Dr Bowman has witnessed the birth of the universe. The absurd imagery, along with the eerie classical music, arguably classifies the sequence as both terrifying and intellectually stimulating.

After Dr Bowman has witnessed himself ageing and the fourth and last monolith appears at the foot of the bed, and when the camera cuts back to Dr Bowman, he is now the star child. However, instead of the now well-known leitmotif of Ligeti's *Requiem*, the spectator now hears the first humming tunes of Strauss' *Thus Spoke Zarathustra*. As the camera performs a dolly shot towards the monolith. In accordance with the camera moving closer towards the monolith, the music intensifies before it cuts back to outer space, where Dr Bowman now floats as the star child, looking down on the Earth, before he turns to the camera and looks the

spectator directly in the eyes. From this arguably bizarre scene, one might be able to suggest that the fourth monolith represented the next step in human evolution since the same Strauss piece was also heard when the primates evolved to become human beings. Furthermore, the same Strauss piece was also used in the title sequence of the film, in which it was utilised to represent something sublime. Therefore, based upon this observation, it is possible to argue that the monolith itself is a sublime object. Throughout the film, the unnerving leitmotif of the monolith has made the spectator feel anxious about what the monolith was capable of, but by using *Thus Spoke Zarathustra*, Freeland might suggest that the spectator is forced into a new mental situation, that is cognition or thought, thanks to which the spectator categorizes and labels the sublime object that thus becomes elevating and pleasurable.

To end the analysis of elements of the sublime within the film, it is worth noting that by ending the film in outer space with Richard Strauss' *Thus Spoke Zarathustra*, it is possible to argue that the film ends with a scene that can be classified as sublime. Hereby suggesting that the opening and the end of the film 'wraps' the whole narrative in a feeling of pain, pleasure and elevation. Therefore, based on the analysis, this thesis suggests that Stanley Kubrick's *2001: A Space Odyssey* is perceivable as an " artwork of aesthetic praise, signifying that the work has a grandeur or a superlative kind of greatness" (Freeland in Plantinga 1995, p. 67),

THE STRUCTURE OF SYMPATHY IN 2001: A SPACE ODYSSEY

When initiating an analysis of the spectator's engagement with the characters in *2001: A Space Odyssey*, one might at first glance find it rather difficult to apply Murray Smith's structure of sympathy to the characters. Even though the film has a runtime of 142 minutes, it merely contains 50 minutes of dialogue. Furthermore, the main characters of the narrative changes with each act, which might make the spectator cautious about investing too much sympathy in the

characters. However, it is not impossible, therefore this section will analyse the characters of the narrative and investigate whether or not the spectator will be able to find him- or herself in a sympathetic allegiance with the artificially intelligent computer, Hal.

THE PRIMATES

One might argue that the spectator faces a challenge in distinguishing these primates from one another, as they look somewhat similar. Thus, suggesting that, in terms of Murray Smith's 'person schema' it will initially be difficult for the spectator to place them in a social role, despite the fact that these primates arguably possess a discrete human(oid) body. However, because they look so similar to each other, they prove difficult to individuate. Nevertheless, an attempt will be made to utilise Smith's structure of sympathy on the primates, for the sake of the argument that the spectator engages in their existing knowledge of social roles and interaction between fellow human beings. An interpretation of the actions in 'The Dawn of Man' will arguably not be possible if the spectator does not transfer their existing knowledge of social roles onto the primates. However, it is essential to note that interpretations of these primates are subjective and might vary from one individual to another. Nevertheless, for the sake of generalisation, the analysis and interpretation will be built upon a western worldview of social roles and interactions.

When the primates are first introduced, the spectator might recognise them as apes, herbivores, potential prey. The primates appear to be living peacefully alongside the tapirs. At this point, it is still not possible to determine the individual primate's social role. When the primates are attacked by a rival tribe, the spectator might be inclined to feel bad for them, seeing as the new pack of primates appear to be violent. Hereby suggesting that the spectator is aligning with the "morally preferable" characters. It is not until the monolith appears, that the spectator is able to apply social roles to the primates. As the primate 'Moon-

Watcher' discovers the ability to use bones as weapons, the spectator might arguably assign a leader role to him. The notion that he is in a leader role is furtherly emphasised when he leads the attack on the rival tribe, to take back their home.

However, it is arguably not possible to align in a sympathetic allegiance with the primates, since the spectator does not gain any insight to their moral values or thoughts.

DR HAYWOOD R. FLOYD

After approximately 26 minutes of no dialogue, the spectator is introduced to Dr Haywood R. Floyd, a character that, according to Smith, should fit right within the 'person schema'. During this introduction of the character, the spectator will engage in what Smith calls recognition, as the spectator perceives Dr Floyd as a man, middle-aged, average body type. Furthermore, the spectator is told his name and that he is American when he is going through the voiceprint identification booth. Thus, finishing the initial face of Smith's structure of sympathy. However, one might say that the alignment with Dr Floyd has already begun as he began to speak since the recognition mainly revolves physical traits and assumptions on the spectator's behalf. However, for the sake of clarity, the alignment with Dr Floyd will be discussed in the next paragraph.

As mentioned in the theory section of this thesis, alignment with a character is based on the relations the character has with other characters within the narrative. From the first scene, when Dr Floyd enters the space station, he interacts with an elevator operator and the woman at the front desk. From these interactions, the spectator may get the impression that Dr Floyd is polite in his interactions with the personnel. Moreover, he appears to be well-known among the people on the station since the lady at the front desk greets him by stating that it has been a long time since she has seen him at the space station. Moreover, the spectator learns from a phone call that Dr Floyd is also a father and a husband.

As Dr Floyd speaks with some Russian colleagues, he is still acting in a polite manner. However, when the Russians asks about the discovery that has been made on the Moon, he appears to become secretive, which might suggest that the alignment with him might be challenged, since he is restricts the flow of information revolving the character. During a council meeting about the discovery of the monolith. The council expresses concern about how much longer they must make up stories about an epidemic outbreak at the Clavius base, Dr Floyd answers that it is still uncertain, as they have to obtain enough data before going public with their discovery. One might argue that this makes Dr Floyd look like a man of bureaucracy and rationality, which in some cases might be seen as a good trait, which, according to Smith, helps the spectator transition into the last level of character engagement: *allegiance*.

Based on the existing knowledge of Dr Floyd, it is possible to argue that the spectator has achieved an allegiance with him. He appears as a man of concern, knowledge and compassion for his field of science, all of which might be perceived as morally good traits by the spectator. However, it is also possible to argue that due to his interactions with the Russians and the vague answer regarding how much longer the secret must be kept, that some spectators might feel that he is somewhat unreliable. Therefore, the allegiance might be challenged or even prevented. The reason that the allegiance might be challenged can be attributed to the access that the spectator has to Dr Floyd. In Smith's own terms, the allegiance depends on the access the spectator has to a character during the alignment, and whether the information that the spectator is provided with is deemed reliable or not. As the spectator does not have subjective access to Dr Floyd, in the sense of knowing his thoughts and feelings, other than what can be read from facial expressions, one might find his need for secrecy unreliable.

The allegiance is ultimately broken when the spectator learns that Dr Haywood has sent the crew of Discovery One on a mission that might endanger, or even kill the entire crew, without informing them about the purpose of the mission.

DR FRANK POOLE

From first glance, the spectator will perceive Dr Poole as a man, middle-aged, athletic, dark-haired. During a televised interview the spectator begins to align with Dr Poole. The spectator learns that he is Dr Bowman's deputy. In a video message from his parents, wishing him a happy birthday, the spectator gains some more knowledge about Dr Poole's relations, he has living parents that seems to care for their son, he is a well-liked member of his local community, he is idolised by the children at the school where Dr Poole's mother teaches. All this information regarding Dr Poole's relation to other characters aides the spectator to make judgements of his character, to decide whether he or she wants to align in allegiance with him.

During the scene where Dr Poole goes EVA, the spectator hears nothing more than the sounds of the spacesuit, and Dr Poole's breath, which might give the spectator some feelings of anxiety through alignment with the character, since it is arguably a terrifying thought to be floating free in outer space, without any control of where one might be drifting. Additionally, the close-up shots of Dr Poole along with the sound of his breath might suggest that the spectator is 'close' to the character, since the sound of breathing can serve as an indicator of Dr Poole's state of mind while being EVA. Moreover, with the growing suspicion about Hal, the spectator might feel anxious about whether he will try to kill Dr Poole or not. This feeling of anxiety might also confirm that the spectator has now aligned in a sympathetic allegiance with Dr Poole, as he appears to be a reliable character when compared to Hal.

Upon learning that Hal displays evil traits, the allegiance with Dr Poole is strengthened furtherly, and despite knowing that Dr Poole and Dr Bowman plan on disconnecting Hal, the spectator is willing to accept the thought of them 'killing' Hal, since it is a necessity to ensure their own survival.

As Dr Poole goes EVA again, the spectator is now feeling anxious on Dr Poole's behalf, since the spectator know that Hal is about to kill him. From the spectator's reaction it is arguable that the spectator emulates Dr Poole's anxiety about being EVA. Furthermore, it is possible to argue that the cutting technique utilised in the scene, along with the sound of Dr Poole's intensifying breath is manipulating the spectator into feeling scared on Dr Poole's behalf.

In the shots where the camera follows Dr Poole flailing around in what is presumably a state of complete panic, there is no sound present at all, underlining the fact that he is now sentenced to die of suffocation in the vast emptiness of outer space. One might argue that if not all, then some of the spectators might engage partially in what Smith calls *central imagining*, as they might share the feeling of pure panic that Dr Poole displays as he flails away with his dying breath since few things are scarier than the thought of floating helplessly through outer space.

DR DAVID 'DAVE' BOWMAN

From the introduction of Dr Bowman, which from physical representation, the spectator recognises as yet another man, middle-aged, dark-haired. From the television interview the spectator learns that he is the mission commander of the mission. During a conversation with Hal, the spectator learns that Dr Bowman likes to draw, thereby indicating that he possesses creative abilities, which, to some spectators, may be a desirable trait which helps him or her in identifying with the character in an *allegiance*. Throughout the film, Dr Bowman displays positive character traits, an example hereof being that he gives Hal the benefit of doubt, when Hal misdiagnoses an AE-35 unit. However, like with Dr Poole, it is relatively easy to align in sympathetic allegiance with the character, since, as the narrative progresses, the spectator learns, as mentioned in the analysis of Dr Poole, that Hal shows signs of being evil. Therefore,

the spectator sees no problem with the thought of Dr Bowman ‘killing’ Hal. Another reason that the spectator might find both Dr Poole and Dr Bowman to be reliable characters can be attributed to the level of access the narrative provides the spectator with. An example being the scene in which Dr Poole and Dr Bowman talks about how they feel about Hal’s miscalculation of the AE-35 unit. The camera is placed inside of the space pod, which might give the spectator a sense of being included in their thought process, hereby making the characters more reliable.



Figure 13: Dr Poole and Dr Bowman discussing their thoughts about shutting down parts of Hal, while Hal is visible through the window, suggesting that he is always watching.

After Hal has killed Dr Poole, Dr Bowman immediately boards another space pod to go retrieve his body, hereby suggesting that Dr Bowman possesses heroic, or at least honourable traits, which strengthens the allegiance with Dr Bowman even more. However, at this point there is arguably no doubt that Dr Bowman is the morally preferable character.

When Bowman returns to Discovery One, he learns that Hal has gone rogue and will not let him enter the ship. Dr Bowman then has to force his way back into the ship, but since he did not bring a space helmet onboard the space pod, he might risk dying when forcing his way back in. As Dr Bowman prepares to ‘launch’ himself into the emergency airlock the camera cuts to a close-up shot of his face, which provides the spectator with an indication his state of mind, which the spectator then might share with him. At this point it is possible to argue that the spectator’s allegiance with Dr Bowman is so strong that he or she might express the

same emotions. As Bowman re-enters Discovery One, the spectator is able to recognise Dr Bowman's facial expression as being 'set on avenging his crewmembers, a feeling that the spectator might share with Dr Bowman.



Figure 14: Close-up tracking shot of Dr Bowman, signifying that he is set on avenging his deceased crewmembers.

In order to furtherly simulate Dr Bowman's feelings, the camera performs a tracking shot with a shaky cam, solidifying that Dr Bowman is in a state of anger and determination. Furthermore, the whole scene is bathed in red light, which emphasises the anger that Dr Bowman feels. As Dr Bowman disconnects Hal's memory the spectator might notice that his facial expression suggests that he is in distress. The spectator might be able to interpret this as a sign that despite of what Hal has done, Dr Bowman takes no pleasure in 'killing' Hal, which arguably makes Dr Bowman appear even more sympathetic.

During the last act, there is arguably no events or reasons that the spectator's allegiance with Dr Bowman might be challenged. Therefore, the spectator will, for the most part, share Dr Bowman's unmistakable state of fear as he enters the stargate. Nevertheless, the camera provides the spectator with POV-shots and reaction shots to ensure that the spectator has no doubt about how Dr Bowman feels (Figure 15 & 16).

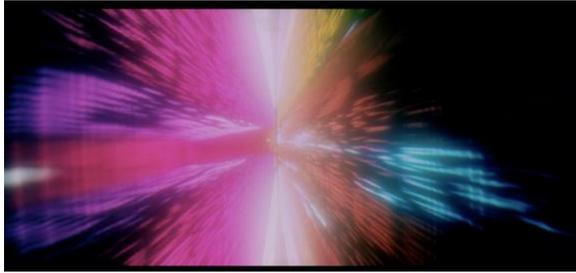


Figure 15: POV-shot of the stargate.

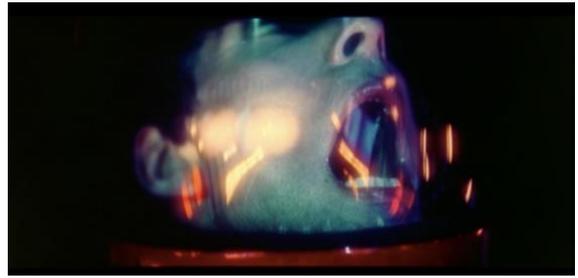


Figure 16: Reaction shot of Dr Bowman, who is visibly terrified

The last scene then takes place in the baroque room, here the camera focuses on Dr Bowman's reaction to the events that just transpired, he is visibly shocked, which the spectator arguably is as well. The camera then cuts back to a POV-shot which elaborates on why Dr Bowman is terrified. In front of him stands an older version of himself, who appears to just as disturbed by the events. One might be able to argue that the reason Dr Bowman is terrified is because he has just seen his doubleganger, which, according to Freud can be a rather terrifying experience, this will be elaborated in another section of this analysis.

The conclusion of the analysis of Dr Bowman's character strongly suggests that despite the narrative is quite sparse in terms of dialogue, he is still perceived as the morally preferable character, since he, unlike Hal, appears to be a reliable character that the spectator is able to identify with in a sympathetic allegiance.

HAL

From the very beginning it is possible to suggest that achieving allegiance with Hal will prove to be rather difficult, seeing as Hal does not meet the most important element of Smith's 'person schema'; a discrete human body with a face. Although Hal might appear to fulfil the rest of the criteria, the lack of a body makes it difficult for the spectator to get an impression of what he might feel. The only thing Hal has is the glowing red 'eye', which one might perceive to be threatening. Therefore, it is possible to argue that Hal is an example of undercut or retarded

recognition of a character, which might challenge or even prevent the possibility of alignment and allegiance.

During the introduction of Hal, in the previously mentioned televised interview, Hal appears to be very confident about his role as the brain and central nervous system of the ship. He praises his own abilities, which the spectator might interpret as either a good character trait. However, it might also suggest that Hal is arrogant, which then challenges the alignment with the character even more. Furthermore, during the interview the characters in the scene are shot at some rather askew angles, which partially can be ascribed to the shape of the spaceship, but it might as well be interpreted as a sign that something is wrong, and since it has been explicitly stated that Hal is the brain of the ship, it might suggest that not everything is as it seems.

Furthermore, Hal appears to be scheming, which is exemplified in the scene where he asks whether Dr Bowman might be having second thoughts about the mission, and that he has heard some rumours about the nature of the mission. The scheming personality is hidden behind Hal's diction, his warm voice makes him sound like he is calm, caring and trustworthy. However, as the narrative progresses Hal challenges the spectator's alignment by misdiagnosing the AE-35 unit and blaming the miscalculation on human error. Hereby, rendering his character as even more unreliable. Another factor that contributes to Hal's threatening character is the way the camera utilises POV-shots from Hal's perspective, the perspective is often placed slightly above eye level of the crewmembers, suggesting that he feels superior to them. Hereby, once again, challenging the spectator's alignment with his character.

In the scene where Dr Bowman and Dr Poole talk, in what they believe to be secrecy, the spectator is provided with another POV-shot from Hal's perspective. In this shot Hal is focusing on their lips, suggesting that he is able to decipher what they are saying.

When Hal kills Dr Poole and the hibernating crewmembers, the possibility of the spectator aligning in sympathetic allegiance with Hal is completely destroyed. Therefore, the spectator, like Dr Bowman, sees it necessary to disconnect Hal. However, as Dr Bowman re-enters the ship and slowly disconnects Hal, there is a slight possibility that the spectator might feel sorry for him, as he sounds genuinely afraid. However, it is possible to argue that this feeling of pity stems from Dr Bowman, who the spectator is sympathetically aligned in allegiance with. Hereby, suggesting that Hal is such an unsympathetic character, that even when he begs for his 'life', the only reason the spectator might feel sorry for him is because Dr Bowman does so.

From this short, analysis of Hal as a character it is possible to conclude that because of the spectator's inability to place Hal within Smith's 'person schema' the recognition, alignment and allegiance is challenged from the very beginning. Moreover, Hal's actions render any chance of alignment in sympathetic allegiance with him as a character, since there are no good moral values to assign to him.

HAL AS A DIGITAL MONSTER

This section will utilise both Michael Szollosy's theory regarding our fear of monsters and projection of the self, and Stephen T. Asma's theory regarding monsterology to investigate which elements of Hal that makes the spectator perceive him as a monster. Moreover, the section will investigate whether the Frankenstein complex is applicable to Hal.

Hal is an artificially intelligent computer from the HAL 9000 series created by humans, which in a figurative sense, can be seen as a way to supersede God, by creating an artificial lifeform. From this perspective one might be able to draw parallels between Hal's creator and Dr Frankenstein, in the sense that despite what noble cause Hal was built for he, like Frankenstein's monster gains autonomy and turns against humankind.

However, with Szollosy's theory in mind, one might also be able to argue that Hal represents humankind's growing search for knowledge which has led us to become hollow shells of cold, unfeeling rationalism, which is exemplified by Hal's adamant focus on fulfilling the mission, that he is willing to kill his crewmembers if they might jeopardise the mission.

Michael Szollosy might argue that Hal is a projection of human rationality that is willing to sacrifice human lives, to make advancements in the name of science. Moreover, Hal can arguably be perceived as a container of anxiety and rationality since Hal represents the human race's projections of violent fantasies, which then makes him become a relentless, destructive, persecuting object. However, since these projections are placed into another entity, the spectator, according to Szollosy, fails to recognise them as their own feelings and instead imagine that they originate from Hal, which causes the spectator to fear Hal.

From Asma's perspective one might argue that what makes us fear Hal is that he reminds the spectator of their own mortality. An example hereof is seen in the scene where Dr Bowman disconnects parts of Hal's memory terminal, in what can be perceived as, in a lack of a better description, a 'digital lobotomy', Hal begs Dr Bowman to stop. In this scene, one might be able to interpret the rapid decrease of Hal's memory as a symbolic parallel to dementia. Hal states that he is afraid and that he can feel his mind is going. As Dr Bowman continues to disconnect parts of Hal's 'brain' his speech slows down and keeps on repeating "I can feel it". Suddenly Hal 'resets' and introduces himself and informs when he became operational and who his creator is. At one point, Hal begins to sing a song he learned in his 'childhood', something that is often seen in dementia patients, as their memory fades away, they return to a state of infancy, and might be able to remember nursery-rhymes or songs from their childhood. As he sings, his voice slows down and eventually fades away, suggesting that Hal has 'died' as a result of Dr Bowman's actions.

To summarise, the Frankenstein complex is indeed present in *2001: A Space Odyssey*. At the time of its release, computers were still a brand-new technology for the common person, as the home computer did not become commercially available before 1977. Therefore, it is possible to argue that people had a fear for this new technology, which might only have been emphasised by Hal's character. Moreover, it is possible to argue that the film is still rather relevant in the year 2020, if not even more so. Artificial intelligence is already implemented in many job sectors, such as farming, customer service, etc. However, when one thinks of artificial intelligence, one might find their mind leading them back to Hal, who has become the posterchild of artificial intelligence gaining true autonomy and causing death and destruction.

THE UNCANNY IN 2001: A SPACE ODYSSEY

Based on Freud's account of the uncanny, it is possible to argue that the spectator might have felt scared of something else than Hal's evil presence, the greatness of the empty space, and the monolith. Once Dr Bowman is has landed in the baroque room, he is visibly terrified of something, which is revealed to be an older version of himself, standing right in front of him staring blankly into the camera (Dr Bowman's eyes). The sight of what Freud has named the double, which stems from the narcissism in the early stages of childhood, which then had a positive meaning since it is perceived as a something, which saves the body from death, like a soul. However, when this narcissism is surmounted, the double no longer has a positive meaning, and turns into an uncanny signal of death (Freud 1974, p, 9).

Based on Murray Smith's theory regarding the spectator's character engagement with Dr Bowman, it is possible to argue that since the spectator is firmly locked in a sympathetic alignment of allegiance with Dr Bowman, the spectator feels the same feeling of uncanniness. I.e., something, or someone, Dr Bowman, who is a well-known character at this point, is

suddenly turned into something unfamiliar, which is unsettling to both Dr Bowman and the Spectator, he is now an uncanny signal of death.

Furthermore, it is arguable that Hal can be perceived as uncanny, since the spectator might raise the question whether he is truly alive or if he, as some of the experts in the film suggests, is merely able to mimic human feelings. By making the spectator question Hal's existence, Jentsch would argue that the story has succeeded in creating an uncanny effect since it:

[...] leave[s] the reader in uncertainty whether a particular figure in the story is a human being or an automaton; and to do it in such a way that his attention is not directly focused upon his uncertainty, so that he may not be urged to go into the matter and clear it up immediately, since that [...] would quickly dissipate the peculiar emotional effect of the thing. (Freud 1974, p. 4).

EX MACHINA

This part of the analysis section will seek out to investigate whether or not the selected theories this thesis has accounted for, are applicable to Alex Garland's 2014 film *Ex Machina*. The analysis will be divided into sections in which relevant theories will be applied to the film.

THE STRUCTURE OF SYMPATHY IN EX MACHINA

This section will, as previously, investigate whether or not Murray Smith's theory about character engagement and 'person schema' is applicable to the characters in *Ex Machina*. Initially, one might argue that *Ex Machina* is more 'viewer friendly', seeing as the spectator is provided with dialogue throughout the entire film. Furthermore, the narrative follows the same

four characters throughout the whole film, whereas *2001: A Space Odyssey* switches out the main characters; Primates, Dr Floyd, the Discovery One crew. Therefore, one might assume that the spectator will find it easier to align in sympathetic allegiance with the character(s) in *Ex Machina*.

Although there are four characters in *Ex Machina*, this thesis will only account for Caleb, Nathan, and Ava since Kyoko can be perceived as more of an occasional presence.

CALEB SMITH

In the introduction of Caleb, the spectator recognises him as a man, tall, slim, red-haired and pale. When Caleb is first introduced to the spectator, he is at work sitting in front of a computer, while sitting in front of the computer, the camera cuts to a close-up shot of the webcam on top of his pc, suggesting that he is being monitored. Moreover, when Caleb picks up his phone, the camera takes on the role of the cellphone, by the use of visual effects, this shot suggests that the cell phone reads his biometrics, in this case, his facial features. Based on this introduction, it is possible to argue that the spectator will engage in Smith's 'person schema' and make assumptions about his character, based on stereotypes. One might get the implication that Caleb is an average Joe, working a 9-5 office job, possibly lonely.

In the alignment phase, the spectator learns that Caleb works as a programmer for the search engine BlueBook, which might give the spectator an indication of Caleb's intellectual capabilities. As the film progresses, the spectator's perception of Caleb being an intelligent man is confirmed through dialogues with Nathan and Ava in which Caleb shows that he is capable of discussing complex scientific theories without any difficulties, which might, at least to some spectators, be a desirable trait that makes them wish to align in sympathetic allegiance with him. During the second meeting with Ava, the spectator learns that Caleb lost his parents in a car accident and he has no significant other. Hereby, suggesting that Caleb is

an honest man, which makes him appear as a reliable character. Furthermore, most of the narrative follows Caleb, which provides the spectator with subjective access to his character. Thus, reinforcing the spectator's perception of Caleb as a reliable character.

As the narrative progresses, the spectator is finding him- or herself aligned in sympathetic allegiance with Caleb, seeing as he is the morally preferable character if one compares him to Nathan, who seems manipulating, intimidating, emotionally unstable, and quite fond of alcoholic beverages. Furthermore, it is possible to argue that because of the spectator's alignment with Caleb, the spectator might find Ava reliable as well, seeing as Caleb senses that there is something off about Nathan, Caleb seems to be able to make good judgements of characters. Therefore, it is possible to argue that when Ava betrays Caleb in the end, it is not only Caleb that feels betrayed, but because the spectator's sympathetic allegiance with Caleb is quite firm, he or she might feel betrayed by Ava's actions, as she leaves him to die.

NATHAN BATEMAN

Before meeting Nathan, the spectator might get a feeling that maybe there is something that seems unreliable about his character. When Caleb gets off the helicopter at the beginning of the film, he is told that the pilot is not allowed to get any closer to Nathan's residence, which might suggest that Nathan is possibly hiding something from the public eye.

When the spectator gets introduced to Nathan, the 'person schema' is once again put to work. The spectator recognises the character as a human and therefore engages in the structure of sympathy. Through the mise-scène of Nathan, the spectator will recognise him as being male, short-haired, bearded, athletic body, glasses. As Caleb and Nathan get introduced to each other, Nathan seems to be friendly and considerate, and he acknowledges that Caleb might be intimidated by standing in front of the founder of the company he works for. However,

soon after meeting each other Nathan begins to show signs of being controlling and manipulative; an example being when he wants to introduce Caleb to Ava, Caleb has to sign a non-disclosure agreement, which Caleb seems sceptical about. Instead of addressing Caleb's concerns, Nathan instead suggests that if he does not sign the papers, he will regret it when his project will be revealed. Nathan piques Caleb's interest, which makes him sign the agreement.

After the first meeting, Caleb and Nathan talk about Caleb's initial thoughts about Ava, during this conversation, one might argue that Nathan shows signs of being narcissistic by misquoting Caleb and calling himself God ("*Narcissistic personality disorder - Symptoms and causes,*" *n.d.*). When Caleb corrects him, Nathan deflects the correction by asking Caleb a new question. It is possible to argue that by showing manipulative and narcissistic character traits, the spectator's alignment with Nathan is getting disturbed. The alignment with Nathan is even more disturbed when Ava tells Caleb that Nathan is not Caleb's friend, and not to trust him. Additionally, when Nathan reacts in an aggressive manner towards his silent servant, Kyoko, it might convince the spectator that despite Nathan's friendly behaviour towards Caleb, Ava might be right about him not being reliable.

When Nathan tries to redeem himself by letting Caleb know how Ava's brain was created, he tells Caleb that he hacked the world's cell phones, computers, web cameras, etc., he is arguably perceived as an even more immoral character. At this time in the narrative, it is possible to argue that as Caleb starts to sense that there is something wrong with Nathan, so does the spectator which strengthens the spectator's allegiance with Caleb.

When Caleb makes Nathan pass out from drinking, Caleb steals Nathan's key card and gains access to Nathan's computer, where he learns that Nathan has been working on prototypes of Ava for quite some time, he learns that the prototypes have either been killed or turned insane by being locked in the room where Ava now resides. As Caleb enters Nathan's

room, he finds the deceased prototypes, which one might argue is a blunt analogy of Nathan hiding skeletons in his closet (figure 17).

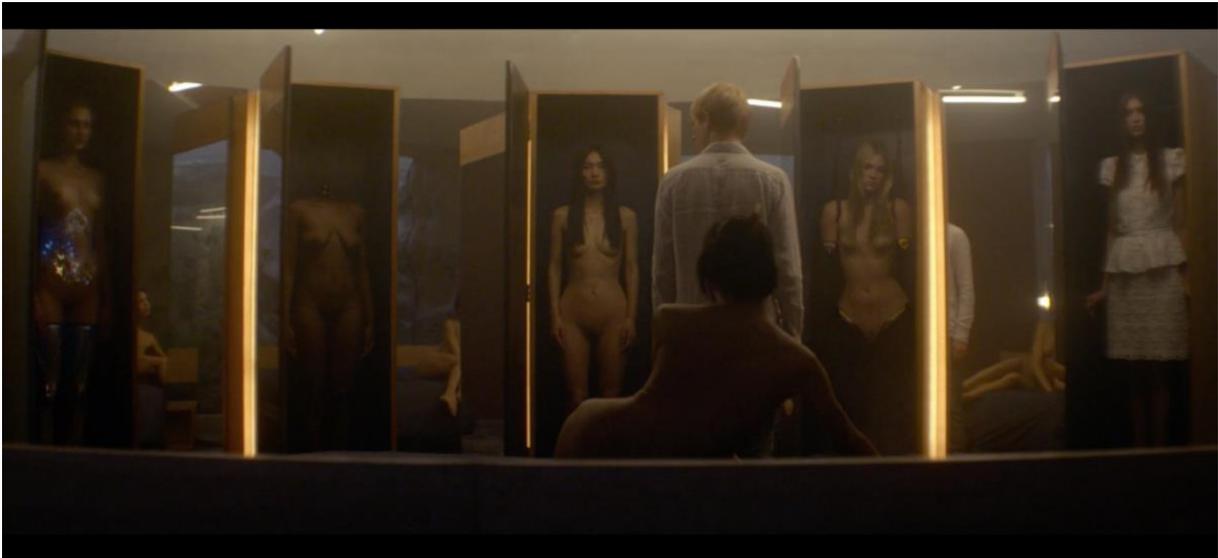


Figure 177: Caleb discovering that Nathan keeps the 'deceased' bodies of previous artificially intelligent robots.

At this point in the narrative, it is arguable that the spectator's alignment with Nathan has been so challenged that it has prevented allegiance, by displaying immoral, narcissistic and manipulative traits, there is no apparent reason why any reasonable spectator would find him- or herself aligned with a sympathetic allegiance with Nathan. Therefore, when he ultimately gets stabbed in the back by Kyoko, and then once more in the chest by Ava, the spectator does not feel sorry for Nathan. Instead, it is possible to argue that the spectator is rooting for Caleb to succeed in helping Ava escaping Nathan's facility to live her life like a real human.

AVA

When Ava is first introduced, she is standing in a low-key-lighting setting, which obscures her facial features, and therefore, challenges the spectator's ability to engage in the recognition level of Murray Smith's structure of sympathy (figure 18). Moreover, when Caleb enters the room, the camera provides the spectator with an extreme close-up shot of a crack in the window

that separates Caleb and Ava, which might give the spectator the impression that Ava might be violent.



Figure 18: Ava emerging from the shadow.

Furthermore, when the camera is moving the spectator's focus from the face, the spectator is encouraged to take notice of Ava's mechanical body. As Ava approaches Caleb the spectator will notice that Ava has a human face, which, according to Murray Smith's 'person schema' is a crucial element of being able to recognise a character and put it in a social context. Although Ava is quite explicitly portrayed as a robot, it is still possible to argue, that because of her face and the shape of her body that the spectator will try to engage in the structure of sympathy and therefore, recognise Ava as a female, feminine body, 'bald'. Once Caleb engages in conversation with Ava, the alignment level is initiated. From the first meeting, the spectator learns that Ava is capable of conversating and reacting with facial expressions, which might give the spectator the impression that Ava is capable of having feelings.

During the sessions in which Caleb is participating in the Turing test with Ava, the spectator begins to align with Ava. In session two, Ava deliberately turns off the power to the whole facility, to warn Caleb about Nathan. She tells Caleb not to trust Nathan, who the spectator already have seen showing undesirable traits that have challenged the alignment with him. Therefore, Ava can be perceived as a reliable character. Furthermore, the spectator gains access to Ava's 'dreams' of going outside the research facility, which suggests that the spectator has subjective access to Ava, hereby, once again reinforcing the alignment with her character.

This access to Ava's dreams are shown through editing techniques, in which the camera puts on a blurry filter with slightly emphasised colours. As mentioned in a previous section, it is possible to argue that the spectator's alignment and eventually allegiance is obtained through Caleb's perception of Ava, seeing as the spectator perceives Caleb as a genuine and reliable character, the spectator might find him or herself aligning in sympathetic allegiance with the character.

Upon learning that Nathan about deleting Ava's memory, which essentially is the same as killing her, Caleb and Ava comes up with a plan for them to escape, it is arguable that the spectator even roots for them to succeed in their escape attempt since Nathan has turned into the villain of the narrative. Therefore, one might find it possible to forgive Ava for killing Nathan, since it is her only chance of surviving.

However, once Ava has successfully have killed Nathan, the sympathetic allegiance is challenged, one might even say entirely demolished as she leaves Caleb for certain death, as he gets locked inside the facility.

With Murray Smith's theory of character engagement in mind, it is possible to argue that although a character might seem reliable, the narrative still has the ability to withhold information about a character. It is hereby suggested that engaging characters, in some cases, can be seen as a manipulative narrative device.

MAKING OF A BEAUTIFUL MONSTER

When watching *Ex Machina* through Michael Szollosy's theoretical lens, the spectator will realise that the tale of Dr Frankenstein and his monster is quite recognisable within this film; Nathan is living secluded from society, tinkering with creating artificial life. Eventually, his creations turn against him, resulting in his mute assistant, Kyoko literally stabbing him in the back, before his magnum opus, Ava, delivers the killing blow. Furthermore, when Nathan

misquotes Caleb and calls himself a God, one might perceive this as a precursor to him becoming a victim of hubris.

Furthermore, as Szollosy mentioned in his paper, the tale of Dr Frankenstein is the tale of us becoming the monsters. Hereby it is suggestable that since Ava has learned all everything she knows about humans and interactions between humans through the internet, i.e. BlueBook. Therefore, one might argue that she is a living projection of all of humanity's violent fantasies, anxieties and fears. She merely acts as she has learned through excessive internet search provided by Bluebook's users.

Moreover, it is arguable that Ava leaving Caleb for dead can be interpreted as her, ensuring her own survival. By using what is, to her at least, rational thought, she concludes that in order to survive in the real world, no one must know that she is a robot. Therefore, she ensures that her only witness is killed.

Lastly, it is possible to argue that Ava and Kyoko can, through the lens of Stephen T. Asma, be perceived as monsters. However, one might suggest that despite Ava being the main character within the narrative, it is Kyoko that awakens terror within the spectator, at least when it comes to portraying the classic monster. Kyoko appears to be human from the first glance, yet when Caleb discovers that Nathan has been making robots before Ava, Kyoko walks up to Caleb and begins to peel off her 'flesh', almost like a zombie removing putrefied flesh



(figure 19). Thus, making herself what Asma would call a morbid reminder of our own mortality.

THE UNCANNY IN EX MACHINA

Ava raises the same question as Hal, whether she is truly ‘alive’ or if she is just a highly functional robot with the ability to mimic emotions and emotional responses. However, when looking at Ava through Murray Smith’s structure of sympathy, the spectator was able to feel that she was a reliable character with genuine feelings. Instead, one might suggest that the truly uncanny element of *Ex Machina* is Nathan’s silent assistant, Kyoko, since she is not showing any emotion throughout the narrative. She is looking human, but not acting like it, which fits within Jentsch’s description of what a narrative must do to create an uncanny effect. Examples of her uncanny behaviour are shown first when she wakes up Caleb without uttering a word, then when she spills something over Caleb and Nathan addresses her in quite a demeaning manner, she does not seem to respond emotionally at all.

Furthermore, it is possible to argue that when Nathan strikes Ava resulting in her arm falling off, and when he afterwards strikes Kyoko, which rips off her jaw might, at least according to Freud, awake the repressed feeling of castration anxiety, the thought of losing a crucial limb (Freud 1974).

Moreover, one might be able to draw parallels between *Ex Machina* and *The Sand-man* (E.T.A. Hoffmann in Freud 1974), seeing as Olympia was able to drive Nathaniel insane, Caleb’s discovery of Kyoko being a robot, causes him that have a mental breakdown, that leads him to cut his wrist open to ensure that he is not a robot himself.

Additionally, one might argue that another reason that *Ex Machina* as a whole might be perceived as uncanny is because it can be perceived as a “double” of real life. I.e., the current debate about how we can be observed through our webcams, cell phones, internet search

history, willingly providing companies with biometric data (fingerprint, facial scans). BlueBook can be perceived as a fictional version of Google and Facebook, who both have been known to experiment with artificial intelligence, where the latter of the two had to shut down the project after two robots created their own language to communicate with each other. Although the robots were not shut down because of their ability to create their own language, it was what the news put in their headlines (Walker 2017), (Kenna 2017), and (Griffin 2017). Thus, confirming Szollosy's statement from the theory section of this thesis regarding that part of our fear of robots stems from the media.

AVA AS A SUBLIME OBJECT

Upon examining *Ex Machina*, one might find that there are elements of the film that can be classified as sublime. Seeing as the thought of creating artificial intelligence that is able to deceive a human being into believing that it is able to feel genuine emotions of affection, only to kill the 'love interest' to gain freedom is arguably somewhat terrifying. Furthermore, the thought of Ava living among humans completely unnoticed might give rise to the question within some spectators, whether this might soon be a reality, or it already is. It is possible to argue that this thought might awake a feeling of 'rapturous terror' within the spectator, i.e., it calls forth a feeling of pain and pleasure; it is painful to see that Caleb gets left for dead, the thought of an artificially intelligent robot already walking among us is terrifying. However, one might, upon further reflection, find it intellectually elevating, to fantasise about the grandeur of human capability to invent marvellous things.

BICENTENNIAL MAN

Chris Columbus' 1999 film *Bicentennial Man* will be the last subject of analysis within this thesis. When watching this film, one might notice that it is quite different from the other

analysis objects, seeing as there is no robots that oppose a threat to humanity within the narrative. Instead it is a family friendly film about a robot who wishes to become a man. Due to the nature of the film, one might argue that Michael Szollosy's theory regarding the Frankenstein complex will not be applicable to this film. However, it is possible to argue that his theory about projections will still be of use. Therefore, this analysis will primarily focus on Murray Smith's structure of sympathy, and Michael Szollosy's theory regarding projections.

THE STRUCTURE OF SYMPATHY

Although *Bicentennial Man* consists of various characters, the main focus lies on the artificially intelligent robot, Andrew. Therefore, the structure of sympathy will mainly be used to determine how the spectator finds him- or herself in sympathetic allegiance with Andrew.

ANDREW

When first introduced to Andrew, one might find it difficult to engage in the recognition level of the structure of sympathy. However, it is not entirely impossible, therefore, the recognition level will be described as best as possible. Firstly, the spectator will recognise Andrew as a robot, male-shaped, seemingly friendly. However, it is important to note that if the spectator was observant during the title, one might have noticed that the leading role in this film is Robin Williams. Therefore, as Murray Smith has argued, it is possible to argue that spectator will make some presumptions about Andrew as a character, based on Robin Williams' star personae. Seeing as Robin Williams is known to be type casted into portraying likeable characters, one might be led to believe that that will also be the case with this film. Hereby, suggesting that some spectators may be predisposed to certain feelings revolving Andrew, which makes the alignment easier.

As Andrew introduces himself to the family the spectator is explicitly ‘told’ to align with Andrew, this is done by the camera performing a POV-shot from Andrew point of view. Thus, suggesting that the spectator will have subjective reliable access to him as a character. Furthermore, when Andrew presents Isaac Asimov’s Three Laws of Robotics to the family, the spectator gains insight into his basic moral values. Hereby, suggesting that Andrew as a character is not able to cause harm to other humans, which arguably might be perceived as a desirable trait, thus, helping the spectator aligning with him.

As the years pass by, Andrew acquires more social skills and has become quite close with the family’s youngest daughter, ‘Little Miss’. Their close friendship is illustrated through a scene in which Little Miss teaches Andrew to play the piano, the camera then zooms in on their hands as they play, before a fading match-cut is performed, which helps emphasise the passing of time, once the camera pulls back out, Little Miss is now young adult.

By now it is possible to argue that the spectator has aligned in sympathetic allegiance with Andrew, the spectator will, according to Murray Smith, have made a moral evaluation of Andrew, in which one might arguably find Andrew to be a character of good moral. Furthermore, it is possible to argue that Andrew is, in the terms of Murray Smith, the morally preferable character within the narrative, since most of the humans he interacts with seem somewhat hostile towards him, despite Andrew not showing any signs of being a threat to any living being.

As Andrew gets his new human appearance, one might argue that the spectator might feel a sensation of happiness on Andrews behalf. Moreover, to emphasise how far Andrew has come in his quest of being accepted as a real man, the camera performs a ‘split-screen shot’ by filming into the mirror to create a visual effect of Andrew’s present appearance, in comparison to his former mechanical shell (figure 20). From this shot, one might argue that

the spectator feels the same level of amazement because of the sympathetic allegiance the spectator has with Andrew.



Figure 20: 'split-screen effect performed with a mirror, emphasising the change of Andrew's physical appearance.

Andrew can arguably be perceived as a sentient being with real feelings and emotions, since he is able to mourn the loss of his loved ones, which makes him an even more likeable character. Therefore, it is suggestable that the spectator is rooting for Andrew in his pursuit of being acknowledged as a real man.

However, it is possible to argue that if one was to apply Murray Smith's structure of sympathy to all of the characters in *Bicentennial Man* it would suggest that almost all of the main characters can be perceived as characters that the spectator is able to align in sympathetic allegiance with. Since there is no apparent villain in the film, it might be perceived as an easy-going, feel-good film.

MONSTERS OF BICENTENNIAL MAN

When watching *Bicentennial Man*, one gets the impression that there are monsters present, neither is there any feeling of looming threat. Therefore, at first glance, *Bicentennial Man* is merely a sci-fi/rom com. However, if one were to take a deeper look into the film, one might be able to argue that Michael Szollosy's theory regarding projection is applicable to Andrew,

albeit not in the sense that the spectator projects their anxieties, violent fantasies and fears. Instead, Andrew can be perceived as a container in which the spectator projects good parts of one's self, so that he or she can identify with the character. Therefore, in the case of *Bicentennial Man*, projective identification through reflection can be argued to provide the basis for feeling empathetic towards Andrew, and by projecting empathy into Andrew as a container, it is suggestable, based on Szollosy's thesis, that the projection of empathy has given Andrew the ability to feel empathetic, which is one of the key qualities that we deem to qualify us as humans.

Furthermore, if one were to look at *Bicentennial Man* through Szollosy's account of the romantic and gothic context, one might be able to agree with the perception that "we ourselves are becoming the fearsome robots, an empty, mechanical shell of cold, unfeeling rationalism" (Szollosy 2014, p. 435). This becomes evident as the humans around Andrew do not recognise his capability of feeling emotions, they do not acknowledge that he deserves love. At several accounts people insist on referring to him as a robot instead of granting him a personal pronoun. Moreover, it is shown when Andrew seeks to be accepted as a human being by the world congress but gets rejected on the cause that humans cannot accept an immortal man, since it arouses too much jealousy.

THE UNCANNY IN BICENTENNIAL MAN

As it has been suggested throughout the analysis of the film, *Bicentennial Man* is more of a family-friendly film when compared to *2001: A Space Odyssey* and *Ex Machina*. Therefore, it is possible to suggest that the film does not awaken a feeling of uncanniness within the spectator, if one is to follow Freud's account hereof. However, one might be able to argue that there are elements that in another context would be perceived as uncanny, but since the general atmosphere of the film is relatively light-hearted it does not awaken the feeling.

However, for the sake of argument, this thesis will suggest that Andrew as a character experiences a feeling of uncanniness when he returns to the Martin household after 20 years of searching for a robot like himself. Upon entering the house, he is met by Little Miss. However, he learns that it is in fact Little Miss's granddaughter, Portia who happens to be identical to the now much older Little Miss, when Andrew recognises Little Miss, he appears to be agitated by the fact that Portia has stolen Little Miss's appearance and directly states that he does not like it.

However, if one is to follow Gray & Wegner, one should by their definition get an unnerving feeling of the uncanny, since they argue that "human beings like the idea of having robots that do things, but not feel things". (Gray & Wegner 2011, p. 129). Nevertheless, as mentioned earlier in this analysis, because of the representation of Andrew as a character, the spectator's allegiance, the underlying soundtrack, and the humour that can be found in *Bicentennial Man*, the unnerving feeling of the uncanny is held at bay.

DISCUSSING ARTIFICIAL INTELLIGENCE: FRIEND, FOE OR SLAVE?

Bicentennial Man might raise the question, which was also raised in the film. If artificially intelligent robots were to have jobs, should they then get paid? To Little Miss, the answer was obvious. However, it is evident that Andrew was met with some resistance, firstly from Richard Martin, who found it natural that he and his family should receive the money from Andrew's work. And then from the bank manager who, although not having an explicit rule against opening a bank account for a machine, was quite resistant to the thought. Based on this example, among others within the film and the analysis, one might be able to suggest that there are some signs of xenophobic behaviour in the way Andrew is treated.

Furthermore, one might argue that with sentient robots, a natural question will arise: what about ethics? When does it become unethical to turn off an artificially intelligent machine? Moreover, how does one even go to determine whether it is truly an artificially intelligent machine? Since there is no complete consensus on the matter yet, it might be challenging to answer the question of whether it is ethical to turn off a seemingly intelligent entity. However, one might argue that in the interest of the human race's survival, if an artificially intelligent entity acts in a threatening manner, like Hal or Ava, one might have to 'kill' it, in the name of humanity.

Although this thesis has primarily been working with narratives in which the robots are perceived as evil or threatening, there are still cases in which the artificially intelligent computers or robots are perceived as friendly companions to humans. Examples being films and television series such as R2-D2, C3PO, K2SO, etc. in the Star Wars franchise, Samantha, *Her* (2013), Bender, *Futurama* (1999-2013), David, *A.I – Artificial Intelligence* (2001), *Interstellar* (2014), Tony Starks computer assistant J.A.R.V.I.S. (Later known as Vision), in the Marvel Cinematic Universe. However, the film industry primarily focuses on narratives in which the robot plays the villain. One might argue that the number of films that

contain artificially intelligent computers or robots is on the rise is because the technology is catching up with the works of fiction, which piques the general public's interest; the thought of a metallic monster that rises to seek revenge on its human enslavers. As Stephen T. Asma suggested, it is the masochistic side of us that are enthralled by these stories.

CONCLUSION

Throughout this thesis it has become evident that it is possible to apply Murray Smith's structure of sympathy to an artificially intelligent robot. However, in order to become aligned in a sympathetic allegiance with these characters, they must have a discrete human body and a face that is able to express feelings. Otherwise, the alignment and the allegiance can be challenged or even prevented entirely, which was proven to be the case with Hal. However, it has also been proven to be possible for a narrative to use engaging characters to manipulate the spectator by making the character appear genuine and reliable, and then suddenly catch the spectator by surprise. Thus, leaving them with a feeling of being betrayed, which was shown through Ava in *Ex Machina*. Lastly, through the analysis of *Bicentennial Man* it was proved that the spectator is able to align in sympathetic allegiance with a robot, and even perceive him as the morally preferable character, when compared to most human characters within the narrative.

Furthermore, through Cynthia A. Freeland's account for the sublime it has become apparent that we enjoy films that evoke the feeling of "rapturous terror" whether it being the grandeur of nature in *2001: A Space Odyssey*, or Nathan Bateman's artistic genius that calls forth the feeling, the spectator is able to transform the painful feeling into something intellectually elevating.

Lastly, by analysing *2001: A Space Odyssey*, *Ex Machina*, and *Bicentennial Man*, it has become evident that in accordance with Szollosy's theory regarding projection, that there is a pattern that can be found within the relatively limited selection of films. A pattern which shows that when we project our fears, anxieties, and violent fantasies into an artificially intelligent container, it turns into a monster that haunts humanity, which was shown by Hal and Ava. However, if we project our good qualities into a container, such as Andrew, it becomes evident that we are the monsters. Seeing the way Andrew and his fellow NDA-114 robots were

treated in the narrative, might awake some unpleasant connotations to how slaves were treated. I.e., making them do dull chores, hard work that no one wanted to do, and stripping them of their free will. Therefore, it is suggested that based on this analysis, the reason artificial intelligence is often cast to play the villainous threat to humanity, is because it helps us renouncing our unwanted parts of the self.

REFERENCES

LITTERATURE

- Asma, S. (2009). *On Monsters: An Unnatural History of Our Worst Fears*. Oxford; Oxford University Press.
- Freud, S., Strachey, J., Freud, A., & Richards, A. (1974). *The standard edition of the complete psychological works of Sigmund Freud*. (pp. 219-252) London: Hogarth Press and the Institute of Psycho-Analysis.
- Gray, K., & Wegner, D. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1), 125-130.
<https://doi.org/10.1016/j.cognition.2012.06.007>
- Griffin, A. (2017). *FACEBOOK'S ARTIFICIAL INTELLIGENCE ROBOTS SHUT DOWN AFTER THEY START TALKING TO EACH OTHER IN THEIR OWN LANGUAGE*. Independent. Retrieved 31 May 2020, from <https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>
- Kenna, S. (2017). *Facebook Shuts Down AI Robot After It Creates Its Own Language*. Huffingtonpost.com.au. Retrieved 31 May 2020, from https://www.huffingtonpost.com.au/2017/08/02/facebook-shuts-down-ai-robot-after-it-creates-its-own-language_a_23058978/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAKE_Cm2hdfvwNQXsvz9qGOvS08KnDzChu4K8UPSabjd39i55NNaaX09qBC5RmtGV_RRR8HoKVj53aQLyBrbc05Rkb-XqXxctmK5izPJbDYqH8aYGhaXKhTZfXMHeZwG-xaP8naa-Lt8fJWmt3S8x-NBNgZ8tKLDaEo-cyEPGVaLY.

- Kivnick, H., & Demiris, G. (2018). ROBOTS AS FRIENDS? A GERONTOLOGICAL EXPLORATION OF FILM, ARTIFICIAL INTELLIGENCE, AND ETHICS. *Innovation In Aging*, 2(suppl_1), 52-52.
<https://doi.org/10.1093/geroni/igy023.194>
- List of artificial intelligence films*. En.wikipedia.org. (2020). Retrieved 1 June 2020, from https://en.wikipedia.org/wiki/List_of_artificial_intelligence_films
- Narcissistic personality disorder - Symptoms and causes*. Mayo Clinic. (2020). Retrieved 1 June 2020, from <https://www.mayoclinic.org/diseases-conditions/narcissistic-personality-disorder/symptoms-causes/syc-20366662>
- Plantinga, C. (1999). *Passionate Views: Film, Cognition, and Emotion*. London; The Johns Hopkins University Press.
- Schneider, S. (2016). Asimov's "Three Laws of Robotics" And Machine Metaethics. *Science fiction and philosophy: from time travel to superintelligence* (2nd ed.). (pp. 290-308) West Sussex, England: Wiley Blackwell
- Smith, M. (1995). *Engaging Characters: Fiction, Emotion, and The Cinema*. Oxford; Oxford University Press.
- Sparkes, M. (2015). *Top scientists call for caution over artificial intelligence*. Telegraph.co.uk. Retrieved 1 June 2020, from <https://www.telegraph.co.uk/technology/news/11342200/Top-scientists-call-for-caution-over-artificial-intelligence.html>
- Szollosy, M. (2017). *Freud, Frankenstein and our fear of robots: projection in our cultural perception of technology*. *AI & SOCIETY*, 32. (pp. 433–439.)
<https://doi.org/10.1007/s00146-016-0654-7>
- Targeted News Service. (March 26, 2018 Monday). *50 Years After Release, '2001' Remains One of the Greatest*. Targeted News Service. Retrieved from <https://advance-lexis->

com.zorac.aub.aau.dk/api/document?collection=news&id=urn:contentItem:5RYR-VH91-DYG2-R121-00000-00&context=1516831

Walker, J. (2017). *Researchers shut down AI that invented its own language.*

Digitaljournal.com. Retrieved 31 May 2020, from

<http://www.digitaljournal.com/tech-and-science/technology/a-step-closer-to-skynet-ai-invents-a-language-humans-can-t-read/article/498142>

FILM

Columbus, C. (1999). *Bicentennial Man* [Film]. Hollywood; Buena Vista Pictures Distribution.

Garland, A. (2014). *Ex Machina* [Film]. Hollywood; Universal Studios.

Jonze, S. (2013). *Her* [Film]. Hollywood; Sony Pictures Releasing.

Kubrick, S. (1968). *2001: A Space Odyssey* [Film]. Hollywood; Metro-Goldwyn-Mayer.

Nolan, C. (2014). *Interstellar* [Film]. Hollywood; Warner Bros. Pictures.

Scott, R. (1982). *Blade Runner* [Film]. Hollywood; Warner Bros.

Spielberg, S. (2001). *A.I.* [Film]. Hollywood; Warner Bros. Pictures, DreamWorks Pictures.

Whedon, J. (2015). *Avengers: Age of Ultron* [Film]. Hollywood, Walt Disney Studios Motion Pictures.