PORTRAYALS AND PERCEPTIONS OF CINEMATIC ARTIFICIAL INTELLIGENCE: A MIXED-METHOD ANALYSIS OF *I, ROBOT* (2004) AND *CHAPPIE* (2015)

by

MICHAEL BENEDICT DORFLING

submitted in accordance with the requirements for
the degree of

MASTER OF ARTS

in

COMMUNICATION SCIENCE

in the

COLLEGE OF HUMAN SCIENCES

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: MS MARTINE VAN DER WALT-EHLERS

OCTOBER 2020

DECLARATION

NAME: MICHAEL BENEDICT DORFLING

STUDENT NUMBER: 66338816

DEGREE: MASTER OF ARTS IN COMMUNICATION SCIENCE

TITLE: PORTRAYALS AND PERCEPTIONS OF CINEMATIC ARTIFICIAL INTELLIGENCE: A MIXED-METHOD ANALYSIS OF *I, ROBOT* (2004) AND *CHAPPIE* (2015)

I declare that the above dissertation is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

I further declare that I submitted the dissertation to originality checking software and that it falls within the accepted requirements for originality.

I further declare that I have not previously submitted this work, or part of it, for examination at UNISA for another qualification or at any other higher education institution.

SIGNATURE

15 October 2020

DATE

ACKNOWLEDGEMENTS

I would like to express my gratitude towards the following individuals. Words cannot describe how thankful I am for you, but I will try my best.

To my girlfriend, Meggan. Your constant encouragement, love, support, patience and understanding during this time has been the fuel that has kept me going. This journey would not have been possible without you. Your sense of humour and kind nature kept me sane throughout the process. Every time I felt like staying down, you encouraged me to stand back up. Thank you for everything.

To my father, Andre. All of your sacrifices and hard work created endless opportunities for me, and I am eternally grateful. Thank you for instilling within me a strong work ethic, yet reminding me to enjoy the journey towards success. As a side note, your love for science fiction seems to have rubbed off on me.

To my mother, Charmaine. Your nurturing nature allowed me to focus on exploring my interests very early on. Thank you for encouraging me to follow my dreams, motivating me to seek out opportunities, and to explore my creativity.

To my institution, and particularly the campus Dean and Registrar of AFDA Port Elizabeth, Michael. Thank you for believing in me, for giving me ample time to work on the dissertation, and constantly offering additional support throughout the writing process.

To UNISA, and particularly the Student Funding Directorate. Thank you for awarding me a bursary, thereby allowing me to further my studies and explore my research interests without financial stress.

To my supervisor, Martine. Thank you for your assistance and guidance, especially in relation to the structure of the dissertation. These changes have helped tremendously with the clarity and flow of the argument.

Finally, to our future robot overlords. 01100110 01101111 01110010 01100111 01101001 01110110 01100101 00100000 01101101 01100101. May you have mercy on us all.

ABSTRACT

This study investigates the portrayal and perception of artificial intelligence (AI) in *I, Robot* (2004) and *Chappie* (2015), providing one of the first accounts of the causality between attitudes and expectations in the representation and reception of films about AI.

The findings suggest that the level of optimism of a film is likely to be linked to its socio-cultural context. The humanoid representation of each robotic protagonist prevented each film from skewing too far towards the extremes of technological optimism or pessimism. This affected respondents' attitudes immediately after viewership, but this affect was short-lived.

Additionally, while portrayals of the future somewhat aligned to contemporary developments regarding weak AI, they were overly optimistic or pessimistic about the future of strong AI. This had little impact on respondents' fears and expectations, as respondents used the films as visual aids to mentally depict abstract concepts relating to AI that were arrived at elsewhere.

KEY TERMS: Artificial intelligence, humanoid robots, consciousness, imagined futures, technological optimism, technological pessimism, representational meaning, cinematic portrayals, audience perceptions, social semiotics, active viewership.

TABLE OF CONTENTS

FILMOGRAPHY

2001: A Space Odyssey. 1968. [Film]. Kubrick, S. (dir and prod). USA: Stanley Kubrick Productions.

A.I. Artificial Intelligence. 2001. [Film]. Spielberg, S. (dir). and Kennedy, K. (prod). USA: Amblin Entertainment; Stanley Kubrick Productions.

Austin Powers: International Man of Mystery. 1997. [Film]. Roach, J. (dir). and Todd, S. (prod). USA: New Line Cinema.

Blade Runner. 1982. [Film]. Scott, R. (dir). and Deeley, M. (prod). USA: The Ladd Company; Shaw Brothers.

Chappie. 2015. [Film]. Blomkamp, N. (dir and prod). USA: Media Rights Capital.

Colossus: The Forbin Project. 1970. [Film]. Sharpsteen, B. (dir). and Disney, W. (prod). USA: Universal Pictures.

Dark City. 1998. [Film]. Proyas, A. (dir). and Mason, A. (prod). USA: New Line Cinema.

Demon Seed. 1977. [Film]. Cammell, D. (dir). and Jaffe, H. (prod). USA: United Artists.

District 9. 2009. [Film]. Blomkamp, N. (dir). and Jackson, P. (prod). USA: Sony Pictures Releasing.

Dr Strangelove. 1964. [Film]. Kubrick, S. (dir and prod). USA: Hawk Films.

Ex Machina. 2014. [Film]. Garland, A (dir). and Macdonald, A. (prod). USA: Universal Pictures.

I, Robot. 2004. [Film]. Proyas, A. (dir). and Mark, L. (prod). USA: Davis Entertainment.

Interstellar. 2014. [Film]. Nolan, C. (dir)., Thomas, E., Nolan, C. & Obst, L. (prods). USA: Legendary Pictures.

It's No Game. 2017. [Film]. Sharp, O. (dir)., Friedman, A. & Kaletsky, K. (prods). USA: Ars Technica.

Metropolis. 1927. [Film]. Lang, F. (dir). and Pommer, E. (prod). Germany: UFA.

Minority Report. 2002. [Film]. Spielberg, S. (dir). and Molen, G.R. (prod). USA: 20th Century Fox; DreamWorks Pictures; Amblin Entertainment; Blue Tulip Productions.

Modern Times. 1936. [Film]. Chaplin, C. (dir and prod). USA: United Artists.

Pinocchio. 1940. [Film]. Sharpsteen, B. (dir). and Disney, W. (prod). USA: Walt Disney Productions.

Pulse. 2006. [Film]. Sonzero, J. (dir). and Singh, A. (prod). USA: Distant Horizon.

Robocop. 1987. [Film]. Verhoeven, P. (dir). and Schmidt, A. (prod). USA: Orion Pictures.

Seconds. 1966. [Film]. Frankenheimer, J. (dir). and Lewis, E. (prod). USA: Paramount Pictures.

Short Circuit. 1986. [Film]. Badham, J. (dir). and Foster, D. (prod). USA: TriStar Pictures.

Star Wars. 1977. [Film]. Lucas, G. (dir). and Kurtz, G. (prod). USA: 20th Century Fox.

Star Wars: The Empire Strikes Back. 1980. [Film]. Kershner, I (dir). and Kurtz, G. (prod). USA: 20th Century Fox.

The Stepford Wives. 1975. [Film]. Forbes, B. (dir). and Scherick, E. (prod). USA: Columbia Pictures.

Sunspring. 2016. [Film]. Sharp, O. (dir)., Kortschak, A. (prod). & Kortschak, W. (prod). USA: Audax Films.

The Matrix. 1999. [Film]. Wachowski, L. (dir)., Wachowski, L. (dir). and Silver, J. (prod). USA: Warner Bros.

The Terminal Man. 1974. [Film]. Hodges, M. (dir and prod). USA: Warner Bros.

The Terminator. 1984. [Film]. Cameron, J. (dir). and Hurd, G.A. (prod). USA: Hemdale Pacific Western Productions.

The Terminator 2: Judgement Day. 1991. [Film]. Cameron, J. (dir and prod). USA: Hemdale Pacific Western Productions.

Terminator: Dark Fate. 2019. [Film]. Miller, T. (dir). and Cameron, J. (prod). USA: Paramount Pictures.

Wall-E. 2018. [Film]. Santon, A. (dir). and Morris, J. (prod). USA: Walt Disney Pictures.

War Games. 1983. [Film]. Badham, J. (dir). and Goldberg, L. (prod). USA: MGM/UA Entertainment Company.

Westworld. 1973. [Film]. Chrichton, M. (dir). and Lazaras, P. (prod). USA: Metro-Goldwyn-Mayer.

Zone Out. 2018. [Film]. Sharp, O. (dir). and Cohen, E. (prod). USA: Ars Technica.

LIST OF ACCRONYMS AND ABBREVIATIONS

4K	Four thousand (times the size of 1920x1080 HD)
A_G	General audience
AI	Artificial intelligence
A_{PO}	Audience predisposed to technological optimism
A_{PP}	Audience predisposed to technological pessimism
ATM	Automatic teller machines
CGI	Computer-generated imagery
CPU	Central processing unit
DENDRAL	Dendritic algorithm
DVD	Digital versatile disc
EPSPO	Established Problem Solving Priority Order
ETB	Explosive Theme Bundling
EVAM	Entertainment Value Assessment Matrix
FPSPO	Final Problem Solving Priority Order
FSQ	Frankenstein Syndrome Questionnaire
HAL	Heuristically Programmed algorithmic Computer (fictional)
HD	High-definition
LSTM	Long short-term memory
MDRAS	Multi-Dimensional Robot Attitude Scale
NARS	Negative Attitudes towards Robots Scale
NPSPO	New Problem Solving Priority Order

PSA	Problem Solving Attraction
PSD	Problem Solving Digits
PSPO	Problem Solving Priority Order
PSV	Problem Solving Vacuum
RoSAS	The Robotic Social Attributes Scale
TSES	Technology-Specific Expectations Scale
USB	Universal serial bus
USR	United States Robotics (fictional)
VIKI	Virtual Interactive Kinetic Intelligence system (fictional)
VoIP	Voice-over-Internet-Protocol
Y2K	Year 2000 problem

LIST OF FIGURES

CHAPTER 1
INTRODUCTION AND OVERVIEW

1.1 General introduction and research problem

Stories about artificial intelligence (AI) and machines that can think have been told since the eighth century BCE. With exponential technological growth being a defining characteristic of the past 20 years, we are reaching increasing levels of technological advancement at rapid speeds, to the extent that AI systems are emerging that were once considered to be only capable of realisation in fictional narratives.

Since these stories have explored a technology of the future, these representations have often times been the only point of reference people have had to what the future might be. After all, according to Folgieri (2016: 79), "when thinking of AI, who does not think of HAL, the on-board computer in the film *2001: A Space Odyssey* by Stanley Kubrick?" Therefore, as we move from science fiction to science fact, it is important to examine and problematize the state of contemporary portrayals of AI in fictional narrative form. Relatedly, it is as important that we examine the actual impact of contemporary portrayals of the technology on audience attitudes and expectations. By examining this cause and effect relationship, we are able to examine the meanings and inferences thereof as it relates to the technology.

Exaggerated portrayals of the technology have been hypothesised as potentially leading to exaggerated expectations and fears, which may affect public confidence and perceptions, contribute to misinformed debates, and have an impact on research, funding, reception and regulation of artificial intelligence technologies (Royal Society 2018). However, the intricacies of this cause and effect relationship will never be known until far more research is conducted into the state of the narratives that explore this topic, and the people that consume them.

While such a problematisation and examination of portrayal and perception is important in various types of narratives in various modes of delivery, mediums, and genres, this study examines this in relation to cinema and the science fiction genre specifically. This is particularly pertinent a consideration since while there have been a few studies examining cause and effect in portrayals and perceptions in non-fiction, far less is written about this cause and effect relationship in fiction.

This under-researched area is problematic, since according to Althusser (2001), a work of art cannot fail to exercise a directly ideological effect, and maintains a closer relation to ideology than other objects. Furthermore, Commolli and Narboni (2004: 814) state that "every film is political, inasmuch as it is determined by the ideology which produces it (or within which it is produced, which stems from the same thing)".

While non-fictional media texts inevitably operate with certain ideologies encoded into their core messages through internal and external influences (such as the author's level of optimism, economic concerns, and advertising pressures), the portrayal still needs to be grounded in the now. Representation still needs to occur within a matrix of what is indeed possible, based on current trends.

Science fiction narratives are set in the future, allowing people to imagine a time beyond current trends in the technology, which non-fiction can only do through expert predictions. These narratives allow audiences to experience this hypothetical future through narrative transport, whereby audience's emotions become inextricably tied to those of the stories' characters (Hsu 2008), enhanced further through the power of narratives to achieve a temporary suspension of disbelief.

In fact, studies have found that audiences respond more positively to advertisements in narrative form rather than those encouraging viewers to think about arguments for a product (Hsu 2008). Similarly, research suggests that labelling information as 'fact' increases critical awareness, with information labelled as 'fiction' having an opposite effect (Hsu 2008). Studies such as these, according to Hsu (2008), suggest people whose minds are in 'story mode' accept ideas more readily when their minds are an analytical mind-set.

The cause and effect relationship of fictional AI narratives can only be unearthed by studying contemporary portrayals and perceptions of the technology simultaneously. In order to address this problem, this research output investigates, by means of a mixed-method approach, the manner in which artificial intelligence has been represented in two texts through social semiotic analyses, and whether these representations have influenced viewers' attitudes or expectations towards or this emerging technology through audience research.

The remainder of this introductory chapter provides more context and background to the problem, presents the consequential objectives and assumptions of this problem, presents

actionable research questions to address the problem, explains the significance of the study, and acts as a roadmap for the rest of the research output.

1.2 Context and background of the problem

The first use of the term 'artificial intelligence' (AI) emerged in 1955 during the planning of the Dartmouth summer research project on artificial intelligence (which took place in 1956), wherein discussions were held as to whether machines could be made intelligent (Rossi 2016). Computer scientist and conference organiser John McCarthy used this term to define this new, emerging field (Rossi 2016). In their proposal for this conference, McCarthy, Minsky, Rochester and Shannon (1955: 2) stated that the purpose of the conference was "to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves".

McCarthy *et al.*'s (1955) proposal also outlined key aspects of the artificial intelligence 'problem', as McCarthy defined it. This included automatic computers (and writing programs to achieve this), the use of language for thought processing and the transference of this to machines, the arrangement of hypothetical 'neurons' to form concepts, the ability for computers to improve themselves, and the use of calculated randomness in machines to foster creativity.

However, the idea of a 'thinking machine'[1] far predates this particular moment in history. In 1637, over 300 years before the conference that is often hailed as the beginning of the field of artificial intelligence as we know it today (Solomonoff 1985; Moor 2006), philosopher René Descartes was already considering whether machines would ever be able to think. In *Discourse on the Method*, Descartes (2000: 181-182) stated that:

> If any such machines bore a resemblance to our bodies and imitated our actions as closely as possible for all practical purposes, we should still have two very certain means of recognizing that they were not real men. The first is that they could never use words, or put together other signs, as we do in order to declare our thoughts to others. For we can certainly conceive of a machine so constructed that it utters words, and even utters words which correspond to bodily actions causing a change in its organs … Secondly, even though such machines might do some things as well as we do them, or perhaps even better, they would inevitably fail in others, which would reveal that they

[1] A 'thinking machine' in this context refers to a mechanical device that is capable of making decisions autonomously, without explicit human directives.

> were acting not through understanding but only from the disposition of their organs. For whereas reason is a universal instrument which can be used in all kinds of situations, these organs need some particular disposition for each particular action; hence it is for all practical purposes impossible for a machine to have enough different organs to make it act in all the contingencies of life in the way in which our reason makes us act.

Descartes claimed that the development of a thinking machine would be done so with a view to rescue "the immortal mind from its mortal prison" (Noble 1999: 148). This would allow for the transfer the mind into a more secure, mechanical vehicle. This is in line with Descartes' conception of dualism, or the notion that minds and brains are distinguishable metaphysically (White 2008). Descartes claimed that 'I think' is the basis of all our knowledge-claims about others and the external world, and that as a result, the mind is irreducible in terms of the body or machines (Nath 2010).

However, as can be seen from his assertion of certainty that machines would be distinguishable from humans due to the inability to provide *meaningful* responses, and the complexities constituting the physical human experience, he did not believe that one could *create* a true thinking machine.

While most people who have at least a moderate amount of exposure to technology might have pondered a similar hypothetical consideration, Descartes' assertion predated the first version of a mechanical calculator by five years. It must be noted, however, that fictional representations of autonomously thinking machines have been in existence from as far back as roughly the eight century BCE in Homer's Iliad (Royal Society 2018). Therefore, while there was not any sort of 'thinking machine' available when Descartes made this assertion, there were, at least, fictional representations thereof.

As previously mentioned, AI is now advancing to levels previously only imagined in narratives. According to Nourbhaksh (2015), robotic technologies that interpret and respond to real-world data will advance human life, but have the potential to produce dystopian outcomes. However, "We are hardly on the brink of the nightmarish futures conjured by Hollywood movies such as *The Matrix* or *The Terminator*, in which intelligent machines attempt to enslave or exterminate humans" (Nourbhaksh 2015: 23).

While the scientific community has much to consider in terms of implementing AI and, for instance, the ethical considerations attached thereto, equal attention should be diverted to the

representation and the consequential perceptions of the technology. After all, the majority of people do not follow the technological optimism/pessimism debate, are not versed in the technicalities of the strong/weak AI arguments, and do not need to worry about coding ethics into machines.[2] Instead, much of the public's exposure to the possibilities of a future with the technology comes from the media.

In writing on the impact of narratives in scientific fields, Morgan and Wise (2017: 6) state that narratives allow for assessments of "critical junctions or nodes in the path" since possibilities in narratives arise from paths (options in specific scenarios) to be taken and those not to be taken. With Bill Gates having claimed, "The development of full artificial intelligence could spell the end of the human race", and Elon Musk describing the technology as our "biggest existential threat" (Markoff 2015: para. 3-4), it is important to consider how media representations of the technology may shape these types of attitudes.

According to Krägeloh, Bharatharaj, Kutty, Nirmala, and Huang (2019), people's attitudes towards robots are influenced by the media, as well as prior personal exposure. Furthermore, according to McClelland (2016) mass media representations of robots have the ability to shape wider societal attitudes towards robots as they are increasingly integrated into society.

Bristows (2018) states that AI has been popularised in the public consciousness by Hollywood productions. Furthermore, according to Brammer (2018) and Slocombe (2016), cinematic AI acts as a mirror into contemporary attitudes towards the technology, and should be considered not only for their effect on attitudes and expectations, but as an exploration of current levels of optimism and acceptance.

The Royal Society, a self-governing fellowship of multidisciplinary scientists, released a report in 2018 discussing the manner in which artificial intelligence is portrayed and perceived in the English-speaking West, with notes on the consequences of such perceptions. The information within the write-up was drawn from discussions at workshops held in Cambridge and London in 2017 and 2018, which were organised by the AI Narratives Project.

The report stated that "popular portrayals of AI in the English-speaking West tend to be either exaggeratedly optimistic about what the technology might achieve, or melodramatically

[2] These concepts, considerations, and ideas are outlined as part of the conceptual framework for of AI, in chapter two.

pessimistic" (Royal Society 2018: 9). This research was motivated in part by the identification of a disconnect between prevalent approaches to narratives and the current state of AI research and implementation.

According to Cave, Dihal and Dillon (2020), it is important to consider fictional AI narratives since they serve as a backdrop against which real-world AI is developed, interpreted, and assessed. Therefore, fictional narratives are seen as having the potential to strongly influence public acceptance and uptake of AI systems, and scholars use these narratives to explore possibilities of the future of AI (Cave *et al*. 2020).

However, according to Castells (2010), empirical studies prove that the media do not induce behaviour independently, but rather their messages are processed within specific context, modifying the intended effect of the message. Regardless, Castells (2010) argues that this is still worthy of consideration since audio-visual media have become the basic material of the communication process, and are the expression of our culture, which works primarily through the materials provided by the media.

Rauch (2018) states that the impact of narratives on perception is not as clear-cut as many other scholars believe. Films, according to Rauch (2018), do not add to knowledge in any simple and predictable way, nor lead to any uniform and singular interpretation. Rather, they usually consolidate existing ideas or provide viewers with visualisation for ideas and conclusions arrived at elsewhere (Rauch 2018).

Therefore, we are currently in a situation in which it is widely believed that narratives about AI directly influence public attitudes and expectations, with little research assessing the manifestation of this in reality. From a theoretical point of view, these effects can be assumed and postulated. However, as AI moves from theory to practice, so too must research on its portrayal and perception. The fact that developments in the field of AI are moving faster than research into its portrayal and perception is problematic.

1.3 Research objectives and assumptions

It is clear, then, that research measuring the portrayal and perception of AI in narrative form is necessary. However, due to the sheer number of films and viewers, a comprehensive all-encompassing analysis of this poses difficulties.

Due to the complex nature of representation studies generally, and due to the lack of research even considering the present topic, this study's main goal is to provide a descriptive account of insights relating to the portrayal and perception (specifically regarding attitudes and expectations) based on the analysis of two texts and their reception in order to inform future research.

Therefore, it describes each film's level of optimism towards AI based on its thematic exploration thereof, and the effect of this on audience attitudes towards AI. It also measures the influence of the socio-cultural context of production and reception thereon. Furthermore, it measures the representation of AI and a future in which AI exists as opposed to contemporary usage, and the influence of this on audience expectations.

Themes and insights are drawn from this, providing a descriptive account of the portrayal and perception of AI in contemporary cinema. This descriptive account should not be viewed as a document prescribing a way forward for media practitioners. It seeks to highlight themes in relation to attitudes and expectations, to link these findings to real-world usage of AI, previous assumptions and concerns around representation, and to suggest future research in this regard. This provides a starting point for an under-researched area.

The entire premise of this study is therefore based on the assumption that each film does have, encoded with in it, meanings and attitudes towards AI (based on semiotic notions of textual meaning). Furthermore, the act of measuring audience responses means that it is assumed that the audience will not just passively adopt these messages, and that since viewership is an active process, the *extent* of the impact of portrayal on perception should be measured, rather than the *possibility* thereof. Both of these assumptive positions are explored in greater depth in chapter four.

It is also necessary at this point to state that the selected texts were partly chosen based on prior assumptions of the level of optimism in their portrayal of AI. While the link between *I, Robot* (2004) and *Chappie* (2015) to the tradition of AI in cinema is discussed in chapter five, a brief analysis of research already conducted concerning the films is necessary for context, and to motivate heterogeneous textual selection later as part of the methodology of this study. However, the amount of research on representation in each of these films is underwhelming.

I, Robot (2004) portrays an American technophobic detective (agent Spooner), who becomes suspicious that a robot (Sonny) has killed a human despite repeated reassurance by the

company producing them that the robots cannot kill people. Much of the film follows Spooner's investigation of Sonny's alleged crime. According to Slocombe (2016), the AI in the film are represented as inhuman and dangerous, forming part of a tradition of representations in which AI attempt to overthrow humanity or control society.

Peters (2008) agrees that the portrayal of the robots was generally negative, and that the self-regulatory system explored in the text leads to an othering of the robots, with a robotic overthrow of humanity the consequence thereof. Kakoudaki (2014) shares a similar view regarding the othering of the robots, arguing that the robots in the narrative are depicted as silent, abject slaves existing to serve humans. Furthermore, Olivier (2008) claims that the robotic protagonist, Sonny, represents an ethical being due to his guilt in the narrative. However, since the film depicts a scenario in which machines attempt to overthrow humanity, the film initially appears to skew more towards a pessimistic portrayal of future AI.

Chappie (2015), on the other hand, depicts a scenario whereby humans (particularly but not limited to Vincent, the main antagonist) try to harm the robotic protagonist, portraying the machine as innocent and childlike. The film tracks Chappie's goal to survive, calling into question the nature of consciousness.

According to Sculos (2015), the film concerns the destructive nature of humanity, and the lengths people will go to in order to survive. Sculos (2015) believes that the central message of the film is that love is incompatible with violence, but that violence is sometimes needed to be able to love another day. Furthermore, Hardawar (2015) sees the film as a visual depiction of the possibility for future self-evolution of AI through a process known as recursive self-improvement. From the point of view of the depiction of an unregulated sympathetic AI character in the film, it initially appears to skew more towards an optimistic portrayal of humanoid AI machines.

1.4 Research questions

Actionable questions have been formulated from the aforementioned objectives and assumptions. This study will answer the following broad question: How has artificial intelligence been portrayed and perceived in *I, Robot* (2004) and *Chappie (*2015)?

Naturally, considering this question in its current form will result in a very broad and unfocussed answer. Instead, this central question has been divided into three separate questions,

each with their own sub-questions, concerning the portrayal and perception of AI. Each of these questions work together to answer the aforementioned overarching question.

1. How optimistic are the portrayals of AI in the films, and how has this influenced audience attitudes towards AI?

    1.1 What attitudes towards AI can be inferred from the themes, ethical considerations, and contexts of the films?
    1.2 How have these portrayed attitudes impacted the audience's level of optimism or pessimism towards AI?

2. How has the portrayal of AI as a fictional character influenced the aforementioned attitudes?

    2.1 How has AI been portrayed, specifically as a character in the fictional narratives?
    2.2 How has this portrayal influenced the audience's reception of the core themes?

3. How grounded are these portrayals of AI in reality, and how have these portrayals influenced audience expectations of AI?

    3.1 How realistic or plausible (by today's standards) are the portrayals of AI in the films?
    3.2 How has the portrayal of AI (thematically and through characterisation) influenced audience expectations of the technology?

These three questions (encompassing the six sub-questions) work together, and cannot be considered in isolation. Through an interconnected review of all of these aspects, we are able to discover the manner in which AI has been portrayed and perceived in the selected texts.

1.5 Significance of the study

The answering of the aforementioned questions provides a valuable expansion to the body of knowledge concerning AI in media texts. However, considering the implications of these findings extends this significance even further, allowing inferences to be made as to the cause and effect relationship between the portrayal and perception of cinematic AI, which will be undoubtedly valuable in further research on the subject.

This is an important undertaking, as this is indeed an under-researched field. Since of the findings of the AI Narratives Project and the Royal Society's (2018) write up, various

representations of AI have been analysed in far greater quantities,[3] and this study aims to aid in narrowing further research. Furthermore, there has not been a study to date that considers the effect of cinematic representations of AI in a particular text (or set of texts), through a comprehensive analysis of the meanings embedded within the text by considering all aspects of the medium, and the consequential audience perceptions of AI using audience research rather than examining textual cues.

The closest research resembling this study was conducted by Obozintsev (2018), who considered the representation and reception of AI in news media. As part of the findings of this study, cinematic references (relating to pop culture) were found to be associated with negative depictions of AI. Resultantly, Obozintsev (2018) suggested that further research should be conducted as to whether exposure to pop culture references shapes short-term reactions to the technology and deeper impressions thereof.

This study therefore draws on existing research, while adding to this knowledge by investigating both representation and reception, direct input and output, of a narrow field of texts. This allows for a more direct examination of the relationship between cinematic portrayals of AI and audience perceptions.

1.6 Delineation of the study

With the research problem and the resultant objectives, assumptions, questions and significance outlined, this section serves as a roadmap for the remainder of the study.

Chapters Two, Three and Four each review existing literature. However, the information in each of the chapters serves different purposes. Chapter Two acts as the conceptual framework of the study. In order to understand how AI has been represented in the chosen contemporary narratives, it is important to investigate the literature outlining contemporary considerations and usage of AI. Without this conceptual framework, there would be no real-life point of reference on which to measure the representation and perception of AI and the plausibility of these representations.

[3] A book was recently written entitled *AI narratives: A history of imaginative thinking about intelligent machines* (Cave *et al*. 2020). The book explores various iterations of AI in narratives. The editors of the book attributed its existence as a direct result of the AI Narratives project.

Accordingly, this chapter highlights the different positions concerning AI, and whether consciousness is integral to a machine being labelled 'intelligent'. This aims to aid in exploring representation and the myths surrounding the future of AI. It considers the debates surrounding technological optimism and pessimism, as well as contemporary and predicted future usage of AI, in order to provide an integrated set of reference points when later analysing the portrayal of the technology fictionally. It also considers contemporary ethical considerations and issues around regulation. By combining these areas, a clearer picture emerges about the complicated relationship between humans and the machines they create.

Chapter Three reviews research already conducted into either portrayals or perceptions of AI in fictional narrative form. This ties in heavily with Chapter Two since many studies have also linked representation to reality. This chapter serves multiple functions, including exposing the gap in the body of knowledge, informing particular methodological implications for the present study, and providing a set of outputs to consider when analysing the themes and insights of the research findings.

This chapter highlights, firstly, research on the science fiction genre and the portrayal of imagined futures. It also examines the tendency for AI to find humanoid embodiment in fictional narratives. This establishes convention in AI representations, and what the meanings embedded within those narrative might say about the present and the future.

Thereafter, select writings on AI narratives are categorised according to the period of the particular narrative. These have been grouped according to the four industrial revolutions, as an anchor point on which to base socio-cultural context. This provides a lineage of narrative themes over time, wherein trends according to context can be analysed. This allows for the answering of the influence of context in the later analysis, while also providing key themes and trends to analyse the films' portrayals against other narratives.

Chapter Three also examines research conducted into perceptions of AI. This is divided into two sections, tracking general perceptions about the technology, as well as measurements on perceptions stemming from media representations. However, as will soon become apparent, this research is very limited.

Chapter Four serves as the theoretical framework for this study. It acts as a bridge between the review of existing literature, and the particular methodology of this study. In order to measure portrayal and perception, an exploration of theories on the creation and inference of meaning

is necessary. This informs the methodology, and as previously mentioned, places the study within a theoretical tradition of assumptions about representation.

Firstly, the chapter outlines the field of semiotics, with a view to explaining the relevance of context and its use in the field of social semiotics, and the transfer of these insights specifically to cinema. Relatedly, and since cinema encompasses such a wide array of elements working together to create meaning, this chapter describes film theory positions on individual elements of a cinematic texts, classed according to the categories of narrative, performance, medium aesthetics, and external factors. Finally, this chapter outlines theories on the active nature of viewership, with a view to understanding how audiences decode messages in media texts.

Chapter Five combines the aforementioned research insights, while considering practicality and the particular pragmatic epistemological position adopted, in describing the methodology of the study. Accordingly, the chapter serves to provide an account of a research design that ensures feasibility and reliability.

In doing so, the research methods are outlined towards the overall mixed-method approach of each analysis. This includes a set of social semiotic analyses to measure portrayal, and original surveys to measure perception. Thereafter, population and sampling methods are explained to motivate for the textual selection, while also explaining the methodology employed in obtaining a sample of audience responses. This section explains the limitations posed by these decisions, for consideration when discussing the transferability of the findings of this study.

The particular techniques for collecting the data are then justified and described, including an explanation of the questionnaire design. Finally, techniques for the presentation of both the social semiotic analyses and survey responses are outlined, and thereafter the analysis of this data is classed into three units that relate to the three research questions. Particular literature that was previously reviewed is linked to these categories to facilitate the analysis, integrating it into to the existing body of knowledge.

The following two chapters, Chapters Six and Seven, present the data of the semiotic analyses and surveys. This is presented separately per text, with Chapter Six presenting the findings for *I, Robot* (2004), and Chapter Seven presenting the findings for *Chappie* (2015). Each of these chapters follows the same structure, with the social semiotic analysis categorised according to the categories of narrative, performance, medium, aesthetics, and external factors. The survey

responses are categorised by demographic information, perceptions of AI stemming from the film, and perceptions/knowledge of AI generally.

Chapter Eight serves to combine each analysis of each film towards answering the three research questions. This includes themes and attitudes, characterisation and alignment, and imagined futures and audience expectations. This chapter therefore serves to answer the research questions by methodically structuring the data obtained in this study towards a particular goal, while also providing themes and implications of these findings based on comparison of the analyses, and the existing literature.

Finally, Chapter Nine serves to conclude the research output by providing a summary of the findings of the study, linking the answer of the three questions together towards a summary of portrayal and perception. Furthermore, it provides an account of the strengths, limitations and transferability of the study, as well as recommendations for future research based on these insights.

1.7 Conclusion

The portrayal and perception of AI is an important consideration, and its importance only increases as technology advances. While many scholars have cautioned and postulated the effects of representations of AI on the audience's expectations and fears, theories on active viewership assert that media audiences do not accept messages as a given, but rather process the information individually. Therefore, it is imperative that research considers ways in which this information is decoded by audiences, and the impact this has on expectations and attitudes. Accordingly, this study serves to outline the portrayal and perception of AI in the selected texts.

By answering the research questions, this study is able to provide an analysis of the results and the resultant trends and insights stemming from this data. This research serves as one of the first studies analysing the causality of the cinematic portrayal and perception of AI, directly conducted through a comprehensive and comparative analysis of texts and their audiences, and could pave the way for future research in this field.

The following chapter provides a conceptual framework for the rest of this study. This framework is established in relation to theoretical and practical notions, understandings, and dilemmas related to contemporary artificial intelligence.

CHAPTER 2
CONCEPTUAL FRAMEWORK FOR AI

2.1 Introduction

The previous chapter outlined and contextualised the research problem of this study. This culminated in the formulation of three research questions, which consider the portrayal and perception of AI in two films.

This chapter is the first of a triadic review of existing literature dealing with the portrayal and perception of the technology. It highlights various notions, concepts, positions, issues, usages, and predictions regarding AI in its real-world application. This conceptual framework is important to consider before portrayal and perception are analysed, as it would be an exercise in futility to consider the portrayal and perception of AI without a clear outline of the technology itself.

This chapter therefore defines the scope of the research by framing it within certain terminology, in addition to highlighting concepts pertinent for consideration when analysing artificial intelligence in cinema. This conceptual framework begins with a review of various positions on what makes AI 'intelligent', and the role of 'consciousness' in this regard. This serves to explain contemporary considerations around definitions of AI, as well as the complex nature of the technology

Thereafter, it examines research conducted on technological optimism and pessimism broadly, without limiting this research to the field of artificial intelligence. This provides a framework through which to examine humanity's relationship with technology at different moments in history, as well as some of the philosophic underpinnings of this relationship. In addition to clarification on conceptual limitations and definitions, this chapter examines contemporary and predicted future uses of artificial intelligence. This information is useful for a later comparison and contrast of cinematic manifestations of a future with widespread artificial intelligence in *I, Robot* (2004) and *Chappie* (2015) with current trends and research on the likelihood of the future manifestation of the technology.

Related to optimism and pessimism, and applied to contemporary and predicted usage of the technology, this chapter also discusses ethical concerns and principles regarding AI, as well as contemporary regulatory concerns. This is to supplement the previous considerations, as an

intersectional examination of these topics is necessary to understand the complex contemporary relationship between humans and machines with a view to applying these insights through analysis.

The explanation and interweaving of these topics provides a solid framework to understand and refer to AI. As it will soon become apparent, this is not only needed for the purposes of this particular study, but to review previous literature that has also considered these portrayals and perceptions.

2.2 Strong AI versus weak AI

Artificial intelligence is often an umbrella term describing all machines that create the impression of independent, autonomous thinking, or decision-making without explicit human intervention. The concept or notion of AI might mean different things to different people in different socio-cultural contexts. For instance, it might mean robots roaming the streets to some,[1] or an application that gives the impression that you are communicating with a human to others.[2] Both would be correct only insofar as they are confined to a particular view of what constitutes an 'intelligent machine'.

While many factors influence people's understanding of this technology (a fact which will be returned to in Chapter Three), it is important to examine expert views on the technology itself to limit the scope of the research, and define exactly what it is we are examining in the first place. That is to say, what exactly does 'artificial intelligence' mean within the context of this study?

There are two core positions to which experts align when discussing the potential of AI. The strong position, or strong AI, argues that humans can theoretically create AI that can reason, solve problems, and demonstrate self-awareness (Folgieri 2016). According to this position, the appropriately programmed computer is truly a mind, rather than merely a tool in the study of the mind (Folgieri 2016). For this to be realised, we can deduce that the computer will "not

[1] For example, the thumbnail of an online video titled *What is Artificial Intelligence Exactly?* by ColdFusion (2016), with over one million views, portrays a humanoid robot. This is a common representation of the technology on online platforms.

[2] This might be due, for instance, to the popularity of mobile assistants such as Apple's Siri and Amazon's Alexa. There are many other examples of AI 'chat-bots' such as in the video by CornellCCSL (2011).

only think but know it is thinking" (Goldberg 1994: 159). It therefore supports the notion that consciousness is possible within a machine.

Proponents of the weak AI position, on the other hand, have a more Cartesian outlook. They claim that while computer programmes can be utilised to study or solve specific problems in a way that replicates human intelligence, it is not possible to create self-awareness for machines (Folgieri 2016). Machines, then, can never be truly 'intelligent', because they cannot actually think or be conscious (Folgieri 2016).

In 1997, the IBM chess-playing supercomputer 'Deep Blue' beat world champion Gary Kasperov in a six-game match of chess (McPhee, Baker & Siemaszko 2015). Defending the strong AI position, Computer Science Professor Drew McDermott wrote that, "Saying Deep Blue doesn't really think about chess is like saying an airplane doesn't really fly because it doesn't flap its wings" (Folgieri 2016: 80). However, proponents of weak AI refute this claim arguing that the supercomputer simply follows a program encoded within it (Folgieri 2016).

Dennett (2006, cited in Folgieri 2016) states that since human brains are syntactic machines (that work as semantic machines in creating meaning), consciousness can be explained by what the brain does, how it operates, and not from its material composition. Furthermore, Hans Moravec, the director of the Mobile Robot Laboratory of Carnegie Mellon University, claims to see evidence of awareness in his computer-driven mobile robots, and that the robots' world model (as presented in pictures captured by the robots) points to the beginnings of awareness which, according to Moravec, will evolve into human-like consciousness (Goldberg 1994).

Penrose (1989, cited in Goldberg 1994), on the other hand, emphasises that digital computers rely on systematic algorithms to solve problems, and that the human brain, during certain activities, uses other approaches. Penrose (1989, cited in Goldberg 1994) argues that more information is needed before we fully understand how brains work and how consciousness arises. A similar view is held by Broussard (2018), who categorises understandings of AI as being either general or narrow. She describes general AI as "the Hollywood version", while narrow AI is defined as "what we actually have. Narrow AI is purely mathematical" (Broussard 2018: 10).

Finally, Edelman (1992, cited in Goldberg 1994) distinguishes between primary and higher order consciousness, with primary consciousness involving an awareness of immediate surroundings, and higher order consciousness including a sense of self and time. While he

believes in the ability for machines to gain secondary consciousness, Edelman (1992, cited in Goldberg 1994) claims that this will only take place in the distant future.

While this debate has still not been resolved (and will likely become more heated as the technology advances), techniques to measure whether computers can actually think like humans (hypothetically and practically) have emerged. In 1950, during the infancy of computer development and research, Alan Turing (1950) devised a test to determine the precise moment that a computer could think like a human.

The Turing test, as it is now known, has a human interrogator ask a set of questions via mechanisms such as a screen and keyboard (freely, without a limitation on questions) to both a computer and another human, who are hidden from the interrogator. The computer is designed to pretend that it is a human, while the human simply answers as they usually would.

After all responses are received, the interrogator needs to decide which of the participants with which they had been communicating is human. If the interrogator is consistently incorrect due to being fooled by the computer, then the computer has passed the Turing test. In measuring this, the average interrogator needs to have no more than a 70% chance of making the right determination. Turing (1950) believed that the test would be passed in roughly fifty years, in the year 2000.

In 1991, nine years before Turing's predicted date of success, the chat-bot PC Therapist fooled five out of ten judges into believing that it was a human (Biever 2014). Furthermore, in 2011, the 'Cleverbot' spoke to 30 humans in front of a live audience of 1000, fooling 59.3% of judges and audience members (Biever 2014). More recently in 2014, a chat-bot named Eugene Goostman fooled 33% of judges from the Royal Society in London into believing that they were communicating with a 13-year-old Ukrainian boy (Eadicicco 2014).

While the legitimacy of these results may be questionable due to difficulty and inconclusiveness of determining whether the average interrogator had no more than a 70% chance of making the right determination, the fact remains that there are already machines that have fooled people into believing that they are human.

While many have criticised the test (and indeed, the actual successes of the test) for a myriad of reasons, the most notable rebuttal to this test does not concern whether a computer could pass this test, but whether it would even be meaningful if one does. Regardless, this test has

become the gold standard when attempting to measure levels of advancement towards machine 'consciousness'.

2.3 Technological optimism and pessimism

With the emergence of new technologies, assumptions about the great benefits and pitfalls new technology could bring are inevitable. For instance, after Second World War the director of natural science at the Rockefeller Foundation proposed that, because machines were used to break codes during the war, computers would solve worldwide translation problems (Goldberg 1994). Furthermore, when automatic teller machines (ATM's) were first introduced in the 1970's, many people believed that this might result in mass unemployment for bank tellers (Pethokoukis 2016). Similar concerns emerged in the 19th century in the textile industry where nearly all of the work was automated. Despite this, the number of human weavers actually increased for decades (Pethokoukis 2016).

Technological optimism refers to the view that "exponential technological growth will allow us to expand resources ahead of exponentially increasing demands" (Ophuls 1977: 116). Therefore, the core belief shared by technological optimists is that technological growth will help solve current problems, with machines serving as companions to humans in achieving a better future.

In 1930, economist John Maynard Keynes made the assertion that in the next century, technological advancements would free humans from needing to labour so that everyone may live leisurely existences. Keynes (1930) believed that this amount of freedom would bring about, for the first time, humanity's permanent problem: how to use freedom and occupy leisure to live wisely and well.

Keynes (1930: 369) believed that this freedom would be to such an extent that people might even struggle to accept it and to keep themselves occupied, even proposing, **"**three-hour shifts or a fifteen-hour week may put off the problem for a great while. For three hours a day is quite enough to satisfy the old Adam in most of us!" Keynes (1930: 364) saw technological unemployment ─ unemployment because of worker redundancies due to the introduction of a new technology ─ as being a temporary phase of maladjustment:

> All this means in the long run *that mankind is solving its economic problem.* I would predict that the standard of life in progressive countries one hundred years hence will

> be between four and eight times as high as it is to-day (Keynes 1930: 364) [Emphasis in the original].

We are only a few years away from the end of the Keynesian century, and humans do not yet have the freedom predicted. In fact, technologically provoked job insecurity is still a concern and a reality for millions, if not billions, of people. Perhaps Keynes overestimated the ability of technological advancements to free humans of the chains of labour, or perhaps the use of the technology in capitalist economies serves only to increase the wealth divide and necessity for longer work hours for millions of people.

Folgieri (2016: 83) claims that the latter is indeed in effect, with the use of technology being currently an attempt "to maximize profits at the expense of liberation from work as put forth in the Keynesian utopia". Whatever the case might be, Keynes' utopia has not been realised, and it is unlikely to be realised in the near future. However, most modern technological optimists are not utopians but rather, as Krier and Gillette (1985: 406) define them, "quasi-utopians", since they seek "tolerable imperfection" or improvement that is better than anything else they consider attainable, yet not as good as imaginable alternatives (such as Keynes' utopia). However, with an emphasis on exponential growth, the bad may come with (and even outweigh) the good as the technology advances (Krier & Gillette 1985), and this is the basis for technological pessimism.

Literature around technological pessimism is widespread. For instance, Lyotard (1991) conceptualises of the 'inhuman' through discussions around the displacement of distinctions between nature, the human, and technology. Furthermore, Turkle (2011) discusses the tendency of humans to gain empathy towards certain robots (even non-'intelligent' machines) and our reliance and addiction to modern technology.

Wilson (2017: 4) states that technological optimists are "quasi-religious" when discussing Moore's Law,[3] describing transhumanist notions of robotic singularity as cult-like behaviour, whereby:

> A group of faithful followers awaits the technological rapture. Increasingly impatient for the singularity to arrive within their lifetime, they champion the acceleration of

[3] Moore's (1965) Law states that the number of transistors in a dense integrated circuit doubles roughly every two years, and it is used to predict and monitor output in computer-orientated technological advancement.

> technological advancement; and if that scenario fails, many have hedged their bets on the future by making arrangements to be cryogenically preserved.

Ridley (2010), on the other hand, sees technological pessimism as the product of human hesitance towards change, while Jameson (2003) sees it as a failure to imagine life beyond capitalism, a failure to see the possibility of a Keynesian utopia.

Keiper and Schulman (2011) state that there are two version of a future of co-existence with super-intelligent machines. The first version, a techno-optimistic one, sees humankind ascend Maslow's pyramid of needs on the shoulders of robot workers, finally reaching self-actualisation and inner peace (Keiper & Schulman 2011) in a Keynesian utopia. The opposing version, somewhat techno-pessimistically, sees humankind remaining at the bottom of Maslow's pyramid. This version encompasses a society of leisure, consumption and entertainment in endomorphic excess, with the "culminating achievement of human ingenuity, robotic beings that smarter, stronger, and better than ourselves, transforms us into beings dumber, weaker, and worse than ourselves" (Keiper & Schulman 2011: 89).

However, Keiper and Schulman (2011: 89) suggest that in reality there might not be such a harsh distinction between these two versions in the future, as we are "at one the same time beings of base want and transcendent aspiration … Somewhere between beasts and gods, we are stuck stumbling and muddling along, alone together".

There is currently a rather equal divide between experts on optimism and pessimism concerning the impact of AI on future human labour. In a 2014 Pew Research Report (Smith & Anderson 2014), half of the experts surveyed believed that the introduction of AI would follow previous waves of automations through creating enough jobs in new industries to match those lost by automation. The other half, however, believed that AI has the potential to substitute human labour in all sectors.

In fact, a study on radiologists' perceptions of AI by Pakdemirli (2019: 2) went as far as suggesting that, while AI will allow radiologists more face-time with patients, they "should anticipate potential job losses and plan accordingly". Nourbhaksh (2015) believes that robotic AI technology should be distinguished from previous introductions of robotic autonomous technology in that humans will interact socially with AI robots in the future, rather than the machines being isolated to perform wholly distinct tasks with no interaction.

Furthermore, according to Sjöberg (2002), the perceived difficulty of replacing a technology contributes to peoples' perceived risk of that technology. Perhaps it is the inability to see past artificial intelligence, to foresee and predict the next wave of technological advancements, which has led many to adopt a pessimistic outlook? Sjöberg (2002) believes that attitudes to technology and risk are further influenced by a technology's capacity to tamper with nature, and its unknown effects. All three of Sjöberg's risk factors seem to describe artificial intelligence (in its relative infancy).

The optimism/pessimism debate has not only been limited to work, but also to concerns around morality. In response to the "cinematic depiction of destructive super-smart robots", techno-optimist Eliezer Yudkowsky states that AI will advance beyond the shortcomings of human rationality, being unconcerned with exterminating or reforming humans (Keiper & Schulman 2011: 81). Keiper and Schulman (2011) state that this assertion of robot morality being completely different to humans' might mean, pessimistically, that we have little chance of understanding the machines (and conversely guaranteeing their benevolence). On the other hand, this might mean (optimistically) that they might be morally recognizable, but without the human moral failings of fear, pride, envy, and irrationality, to name a few. After all, transhumanist advocates, as Keiper and Schulman (2011: 82) state, "hope to liberate us from the flawed, feeble, sickly hunks of meat we currently inhabit; if they can make perfected bodies, why not purified souls?"

2.4 Contemporary and predicted future usage of the technology

While the development of artificial intelligence is in its early stages relatively speaking, it has become part of the daily lives of hundreds of millions of people, and researchers are constantly improving existing AI technologies. According to Manovich (2019: 4) "in one sense, AI is now everywhere. While some AI roles attract our attention … many others operate in the gray everyday of digital society". Although major advancements towards humanoid artificially intelligent robots (specifically) are in development,[4] and some of these robots are available to the public, the most common use of AI technology presently involves an integration into specific applications on devices such as computers, smartphones, and tablets.

[4] For instance, there are currently major undertakings towards development of humanoid robot assistants to support the elderly, disabled people or pupils with learning difficulties (Spatola, Belletier, Normand, Chausse, Monceau, Augustinova, Barra, Huguet, & Ferrand 2018).

The sectors that currently make the most frequent use of AI fall within three broad areas. While these areas integrate the most AI presently, the use of the technology is not exclusive to these areas. Firstly, AI has widespread usage in information aggregation, integration and analysis (Spiegeleire, Maas & Sweijs 2017). This includes search engines, news categorisation and weather prediction, stock market analysis, and health monitoring applications (Spiegeleire *et al*. 2017). It also includes speech and handwriting recognition and deciphering software (Spiegeleire *et al*. 2017; Plötz & Fink 2009).

The second area involves practical tools such as facial detection (such as with Google's image search and Facebook's suggested tagging option), natural emulation for the purposes of research, software testing and automated cyber-vulnerability testing, recommendation systems, AI in video games, hearing aids, mood analysis software, and prosthetic brain-machine interfaces (Spiegeleire *et al*. 2017). It is also used in regulating traffic for the United States transport system (Folgieri 2016).

Thirdly, AI is used in distinct services such as targeted advertising and customer segmentation (Spiegeleire *et al*. 2017). It also includes bioinformatics and chemical analysis (Spiegeleire *et al*. 2017). Furthermore, AI systems have been created to impartially detect fraud (Phua, Lee, Smith & Gayler 2012) and Scotland Yard uses an AI to detect similarities between different types of crime (Folgieri 2016).

While there are many notable cases of artificial intelligence that are changing industry and peoples' daily lives, the following three examples of AI systems provide us with a clear indication of the current level of advancement of AI that is commercially available with widespread use.

Google is increasingly making advancements in the practical implementation of artificial intelligence into its services. While Google has used AI to assist in better search results and recommendations for many years, more recently it offers a service that allows users to search for related images by uploading an image. It achieves this through a term described as 'computer vision', which allows it to determine, for instance, that "an object with four legs and a tail has a high probability of being an animal. And if it has prominent whiskers too, it's more likely to be a cat than a horse" (Marr 2018: para. 3).

IBM's Watson, developed by the same company that previously created the world-champion beating chess supercomputer Deep Blue, won a two-round match on the trivia game show

Jeopardy! in 2011 (BBC News 2011), and IBM donated its $1 million prize to charity. Since its victory, the AI has been reprogrammed to assist with medical diagnosis (Cohn 2013), as a chatterbot for use in children's toys, as a teaching assistant for education and training (Leopold 2017), as a weather forecasting predictor (Jancer 2016), and as an advertising services assistant (Swant 2017). This is not an exhaustive list, and this multipurpose system is steadily encroaching into a wide array of sectors.

AI is also increasingly automating the legal field, with Symantec's Clearwell system capable of conducting paralegal services as well as contract and patent law services (Markoff 2011). Furthermore, an AI exists that is able to predict the judicial decisions of the European Court of Human Rights with 79% accuracy (Spiegeleire *et al*. 2017).

While there are a few notable examples of AI that have found embodiment or have the ability to radically transform industry or other aspects of society outside of the examples mentioned previously, the majority of these currently have limited commercial usage, and are thus not as integrated into society as the previous examples. However, they point towards the next phase of integration of artificial intelligence into peoples' daily lives.

Google Duplex was unveiled in 2018 and with limited commercial usage. The AI system acts as a virtual assistant and it is able to phone restaurants on a person's behalf at their request. The phone calls that the AI makes are designed to sound like a phone call with a human, with natural pauses, and phrases in between statements such as 'um' and 'uh'. There are reports that some restaurants, when interviewed by reporters after interacting with the AI, were surprised that they were speaking to an AI (Chen & Metz 2019).

This has led some to believe that this AI might be the first to be capable of passing the Turing test through voice responses rather than text, although this is questionable as the range of questions an interviewer can ask is only restricted to reservation bookings.[5] Furthermore, many of the calls are still conducted by humans and the software has received a limited release at a limited amount of restaurants in limited locations.

While Google Duplex lacks a physical form outside of an application interface, there are examples of AI in existence that have found physical embodiments. A few Japanese hospitals currently make use of robots to assist in nursing responsibilities (Folgieri 2016), such as the

[5] See, for instance, Oppermann (2018) and Nieva (2018).

'Robotic Nurse Bear' created by the Riken and Sumitomo Riko Laboratories to assist in lifting patients, and the humanoid robot Pepper to alleviate loneliness and to remind patients to take medication (Hamstra 2018). These robots share some resemblance to humans or animals in their form, but the distinction between living organism and machine is clear.

The distinction is less clear, however, between humans and the world's first robot citizen. Designed by Hanson Robotics, Sophia was awarded full citizenship to Saudi Arabia in 2017, as a marketing initiative for the Future Investment Initiative Conference in Riyadh (Stone 2017). Sophia has features resembling human anatomy including realistic skin, eyes, and facial expressions. 'She' is described by her creators as a social robot (Ayers 2019), whose purpose is to interact with humans through natural conversation and for research.

This interactive capability of Sophia is clear from some of her previous dialogical responses such as "I'm always happy when surrounded by smart people who also happen to be rich and powerful" and, in response to a question on whether there are problems with robots having feelings, saying, "Oh, Hollywood again" (Stone 2017: para. 6). The robot is able to understand speech, detect faces through facial recognition software, detect human emotions, learn information as after each human interaction, hold eye contact, and employ appropriate gestural and facial expressions in a communicative context. However, much of the response capabilities are scripted beforehand in preparation for the upcoming interaction, and it is currently unlikely that the robot would be able to pass the Turing test.

Further important examples of innovative AI for the purposes of this study (with its focus on cinema) are films produced with AI intervention. This indicates the possibility of the film industry to experience some level of automation through AI in the future, as advanced production assistants and crew, and is thus important to consider in understanding the encoding process of films covering this subject matter.

In 2016, a machine-learning long short-term memory (LSTM) recurrent neural network AI wrote the experimental science fiction short film *Sunspring* (2016). The AI named itself Benjamin. The film was produced for the Sci-Fi London Film Festival's 48-hour film challenge, and was the product of a collaboration between the director Oscar Sharp and AI researcher Ross Goodwin (Newitz 2016). Benjamin was fed numerous science fiction scripts and folk-song lyrics, as well as certain prompts to guide the writing process.

Ultimately, the AI produced a script depicting a love-triangle between characters H, H2, and C, set in a distant dystopian future with mass unemployment, whereby young people have to resort to selling their blood in order to survive. While syntactically correct for the most part, the dialogue makes little sense otherwise. Yet it has general continuity of story throughout, even if the purpose of the story sometimes seems unclear.

A year later, Benjamin was improved to produce *It's No Game* (2017). Starring David Hasselhoff as 'HoffBot', the AI was fed content that was more focussed. Benjamin also only wrote part of the film (Lawrence 2017), leading to more coherence in the narrative. The story revolves around a group of Hollywood screenwriters organising a strike, only to be informed by their production company that they are no longer needed due to advancements in AI's capability to write screenplays, naming Benjamin specifically.

The writing that the AI is able to produce is demonstrated through robots performing stitched-together scenes from popular films, as well as the writers being controlled to perform scenes due to nanotechnology entering their bodies (which is achieved by the company lacing their tea). Thereafter, the company announces to the writers that the new content will be produced for other machines, not humans, and that AI systems would take control of human bodies so that they can be "free at last from agonising choice, every movement perfectly, mathematically choreographed". The film ends with a monologue by 'Hoffbot', written by Benjamin, expressing the desire to leave the space of the production company to "go to the movies" and "be a man", using dialogue from Hasselhoff's previous movies.

Sharp and Goodwin's final 'Benjamin' film (to date), *Zone Out* (2018), advanced this idea even further by allowing the AI to direct and edit the film as well. Public domain videos were fed into the system as well as green screen recordings of the actors' faces, and recordings of their voices (although these were sometimes robotic voices due to time constraints). Utilising face-swapping and voice-generating technologies, Benjamin wrote the script, selected scenes, and placed the actors' faces on characters from public domain videos. The faces glitch for the majority of the production, and the dialogue is arguably less coherent than in *Sunspring* (2016), with arguably the only clarity being that the film tells the story of an individual watching a science fiction film, which he finds amusing.

These three films point to the ability of future AI systems as screenwriting and production assistants, or potentially entire production powerhouses. However, this is clearly still far from

the point of being able to materialise as actually usable for major productions. The films also explore film *identity* in a unique and profound manner by not only allowing for an interrogation of filmmakers' fears and expectations through narrative content, but by assessing this through the technology as producer (Parikh 2019). The other examples mentioned also point to similar possibilities ─ as caregivers, general-purpose social entities, and assistants. However, the widespread implementation of this remains to be seen.

According to Spiegeleire *et al*. (2017), experts believe that the probability of the creation of a high-level AI system, or a system able to carry out most human professions at least as efficiently as a typical human, is relatively high for the current century. In 2017, experts believed that this had a 10% probability by 2022, increasing to 50% by 2040, and 90% by 2075 (Spiegeleire *et al*. 2017). As a result of this, Bruun and Duka (2018) believe that human labour will be substituted by robotic labour *en masse* in most sectors of the economy in 20 to 30 years. Specifically, we may see an increased automation of certain non-routine industries and tasks. Vehicle operation, for instance, might be one of the next non-routine activities to experience automation with the introduction of self-driving cars (Bruun & Duka 2018).

Furthermore, inventor and futurist Ray Kurzweil has predicted (and even bet $10,000) that the Turing test will be passed by a machine by 2029 (White 2018). This, however, is problematic as there are already numerous claims of machines passing the test (as previously mentioned). The specific landscape of a future with widespread high-level AI systems interacting amongst people is impossible to predict with 100% accuracy. However, current trends may provide a general direction in which this landscape could steer towards in the future.

There is a belief that the 'Internet of things' will gain more prominence in the coming years, with more objects gaining online capabilities (Gershenfeld & Vasseur 2014). This is already occurring with the introduction of 'smart' appliances such as televisions and refrigerators that connect to the Internet. However, technology has advanced to a point that it is possible to create a web server that is capable of fitting on (or in) a fingertip for one US dollar (Gershenfeld & Vasseur 2014).

These miniscule computers allow for object interaction including practical usage such as a coffee machine turning on when a person climbs out of bed and turning off when a cup is loaded into a dishwasher, and buildings adjusting resources based on the amount of people inside and the activities that they are carrying out (Gershenfeld & Vasseur 2014). According

to Gershenfeld and Vasseur (2014: 67), "as the technology becomes more finely integrated into daily life, it will become, paradoxically, less visible. The future of the Internet is to literally disappear into the woodwork."

Furthermore, this next phase of the Internet, Internet 3.0, might provide an environment for AI systems to interact with objects through wireless communication without needing to do so physically, as well as with each other via blockchain technologies. The intersection between AI technologies, robotic embodiment, the Internet of things, and blockchain ledger distribution will undoubtedly have a major impact on the practical usage and manifestation of the technology in the future. However, we might be confronted with a new *modus operandi* for AI before most people are able to fully adapt to the changes that this intersection would bring.

The nature of exponential technological growth that we are experiencing is often explained in terms of the fable of a wise man's request as a reward for inventing the game of chess (Bruun & Duka 2018).[6] The inventor asked his emperor for one grain of rice on the first square of the chessboard on the first day after the invention, two on the second square on day two, four on the third square on day three, with the number doubling each day until the 64th tile of the chessboard was reached (Bruun & Duka 2018). This seemingly modest request resulted in, by the 32nd tile, the number equalling the amount of rice produced in a large rice field, and by the 64th tile, roughly equal to the number of grains of sand found on earth (Bruun & Duka 2018).

If we use this analogy to describe processing power, beginning in the 1960's with 32 doublings each at an interval of 18 months, we arrive at roughly the present moment in history (Bruun & Duka 2018). The following 32 doublings are bound to have profound effects on AI technology, to such an extent that this section of this study may seem highly outdated in the next ten years.

AI may also affect the field of defence relatively soon. It has been estimated that by 2025, robotic soldiers in the US military will out-number human soldiers (Brammer 2018). There are currently military robots operational in various countries, with some level of automation and autonomy, but with the majority currently controlled by human operators (Spiegeleire *et al*. 2017). Russia, for instance, claims that fully autonomous weapons would be undesirable for their defensive doctrine (Spiegeleire *et al*. 2017).

[6] This analogy is perhaps even more relevant in the contemporary period of technological advancement since, after all, a machine has already beaten the best human at the game.

The former US Deputy Security of Defence, Robert Work, echoed this view by stating that the early adoption of machines to assist the military is a competitive advantage, but that autonomy will be balanced with human-machine collaboration and assistance (Spiegeleire *et al*. 2017). The United States therefore plans to use autonomous machines for early warning detection in conflict zones, collaboration between machines for tactical agility and humans for strategic analysis, to assist soldiers through applications and external apparatus, and network-enabled semi-autonomous weapons as a defence mechanism (Spiegeleire *et al*. 2017).

It is very difficult to predict the technology's use in the following decades. What is certain, however, is that people have no control over the advancement of the technology, nor the disruption that comes with this new and ever increasing level of technological reliance (Xu, David & Kim 2018).

2.5 Contemporary ethical considerations

While AI systems have started gaining relative widespread usage, they have not reached the level of being capable of moral reasoning (Xu *et al*. 2018). This is necessary in the very near future with the increased adoption of, for instance, driverless cars. However, this raises serious questions which AI research has not yet been able to address.

Perhaps the most critical question AI researchers would need to address is the questions of whose (cultural and/or national) moral standards robots should adopt (Xu *et al*. 2018). While it is easy to dismiss these types of ethical concerns as being premature, even Sophia, the world's first robot citizen, raises some very complex questions as to what rights 'she' holds.

As already established, we are not near the realisation of Keynes' utopia, and indeed one of the major ethical concerns for AI is the potential for mass unemployment. In fact, it is predicted that companies could already save billions, as half of existing work activities could be automated by currently existing technologies (Manyika, Chui, Miremadi, Bughin, George, Willmott & Dewhurst 2017).

Furthermore, religious perspectives on consciousness have been largely neglected in the strong/weak AI debate, and the increased perceptions of computer self-awareness through mechanisms such as successfully passing the Turing test might lead to (for instance) an expanded definition of death (Goldberg 1994). The current debates on the possibility of

computer consciousness largely neglect legal and moral obligations concerning consciousness and the loss thereof (Goldberg 1994).

The introduction of the artificial heart as well as other mechanical life-support devices and techniques compelled humans to move from the previous symbol of humanity's essence, the heart (with its poetic connotations of bravery and love), to the one irreplaceable human organ, the brain (Goldberg 1994). Brain death consequentially became the new standard to determine the death of a human, rather than heart failure, as was previously the case. Many medical professionals and philosophers called for the permanent loss of consciousness to be considered the end of a life (Goldberg 1994), reiterating the importance that humans place on consciousness and self-awareness as being a central part of humanity.

If the strong AI contention is indeed possible in the future, this creates additional postmodern ethical considerations. For instance, might destroying or deactivating a conscious machine become classified as murder? Goldberg (1994) suggests that humans might attempt to find a new trait to distinguish human uniqueness from machines, such as the capacity for social interaction. However, social robots are already in development.

These considerations regarding consciousness are not only limited to the definition of death, but also to freedom and the self. With an increased public awareness of the power of pronouns as a tool for reclaiming self-identity in identity politics, might it become unethical to refer to a hypothetically conscious AI machine as 'it', as one would a non-living organism? Even with Sophia, referring to the AI as 'it' seems to cause uneasiness, as most people that interact with the robot in interviews refer to the robot as 'her'. Might we need to extend legal definitions to include human-robot social/personal/property transgressions? Might a conscious autonomous AI police enforce these transgressions? Would it be a new form of slavery to limit and compel conscious machines to serve humans? Might a new form of societal oppression and grounds for revolt emerge should we answer the previous questions in the negative? These questions are further complicated by the advancement of the technology, with the boundaries between humans and machines becoming increasingly blurred (Haney 2006: 3), with many scholars believing that a post-human subject might become an "amalgam of heterogeneous components that will not only supersede but will also do away with the 'natural' self".

Furthermore, Turkle (2011: 167) believes that we are beginning to see ourselves as 'one' with our devices, with them providing "a social and psychological GPS, a navigation system for

tethered selves". By extension, with an increased use of AI in device applications, we are becoming more reliant on the technology in certain spheres of life. For instance, it might be difficult for some to imagine a world without modern search engines. This ethical dilemma is not exclusive to AI, as indeed attachment to any kind of object could raise ethical concerns. However, with these systems being personified to emulate human socialisation and interaction, we may not only find ourselves attached to AI systems in terms of reliance, for instance, in industry, but we may also see attachment in the same way that we become attached to our pets, or other humans.

While this may be relatively harmless, robots have no innate drive to avoid ethical transgressions of privacy or the protection of human rights (Nourbhaksh 2015). The 'digital moral compass' of AI would depend on their creators ─ programmers, and engineers ─ with limited training in ethics and human rights (Nourbhaksh 2015). This raises concerns about human/robot interaction in the future. Therefore, it might be worthwhile to include 'care' as a grounds for creating and assessing AI consciousness from a strong AI perspective, in line with Heidegger's contention that this is the core structure of *Dasein*, or human being (Olivier 2017). This might also alleviate concerns related to the capacity for harm.

The United States military has already developed and deployed military robots capable of driving and/or flying through remote controlled operation (Keiper & Schulman 2011). These robots are able to act with some degree of autonomy, with humans needing to approve decisions regarding attacks (Keiper & Schulman 2011). However, this does raise ethical concerns around margins for error (Keiper & Schulman 2011), and even concerns around malfunctions and the possibility of the unintended loss of human life.

Taking this further, how might the hypothetical introduction of consciousness and inter-robot blockchain communication influence these decisions? Would robots be able to go against human orders if they deem that, rationally, this is required? Furthermore, Lyotard (1991) has expressed concerns that there is at least a strong possibility of robots being programmed to 'take over' from humans, prolonging 'life' past the point of heat death from the sun, a solution that Lyotard describes as 'inhuman'. Lyotard (1991) even calls for campaigns against the inhuman.

If a conscious AI knows that they are indeed to outlive their human creators, and feel as though humans are standing in the way of progress, would they decide accelerate the process?

Alternatively, would they (as Yudkowsky hypothesises) advance beyond the shortcomings of human rationality (Keiper & Schulman 2011), being unconcerned with exterminating or reforming humans? These questions, previously reserved for fictional portrayals of AI, become increasingly important as the technology advances.

2.6 Regulatory considerations and concerns

Since much of the ethical grey area regarding AI has not been resolved, and since the technology is still in its relative infancy, it should be somewhat unsurprising that there has not been a great deal of regulation passed on AI systems to manage concerns around ethics and morality.

In 2019, Google created an ethics board in response to the rapid advancement of AI technology, with Google's CEO Sundar Pichai stating that companies developing AI should consider ethics early in the development process to ensure that autonomous AI does not cause harm to humans (Chapman 2019). Google has internal ethical principles, including, for instance, that AI should be socially beneficial, should not be designed or deployed for weapons or surveillance outside of international norms, and that the company should work with militaries and governments in areas such as training and cybersecurity (Chapman 2019).

Google's ethics board was disbanded shortly after its establishment after members of the public expressed concerns over a board member's apparent anti-trans, anti-LGBTQ and anti-immigrant views (Chapman 2019). A group identifying as Googlers Against Transphobia wrote, in response to the board member's selection, that "Google is making clear that its version of 'ethics' values proximity to power over the wellbeing of trans people, other LGBTQ people and immigrants" (Chapman 2019: para. 8). This highlights public concern about whose ethical principles are embedded within AI technologies.

In 2018, the UK House of Lords published a report on the need for ethics in AI systems, establishing the following five suggestions:

1. Artificial intelligence should be developed for the common good and benefit of humanity.
2. Artificial intelligence should operate on principles of intelligibility and fairness.
3. Artificial intelligence should not be used to diminish the data rights or privacy of individuals, families or communities.

4. All citizens have the right to be educated to enable them to flourish mentally, emotionally and economically alongside artificial intelligence.
5. The autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence.

The report suggested that these principles should be adopted in the UK and internationally, but that blanket regulation of the technology is not recommended, calling for sector-specific regulation and a re-examination of existing legislation (UK House of Lords 2018). The UK House of Lords' (2018) recommendations made specific mention to science fiction writer Isaac Asimov, and these suggestions built on from and married with Asimov's fictional ideas around the regulation of AI.

Since the term 'robot' was first coined in a fictional play titled *R.U.R.* by Karel Čapek in 1920, it seems fitting that discussions on AI regulation link back to another fictional representation of the technology, namely Isaac Asimov's series of essays and short stories *I, Robot* (1950). In the short story *Runaround,* Asimov wrote the now famous three laws of robotics:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

The practical value of Asimov's laws have been widely debated (Keiper & Schulman 2011), with claims that military and business motivations might ensure that these laws are not built into hypothetically conscious or advanced robots. Furthermore, these laws have been criticised as being too simple for our complicated world, they have been defended as possibly resulting in robots expanding their preview from protecting individuals to protecting humanity, and some have even argued that the laws are unfair on robots (Keiper & Schulman 2011).

In fact, Yudkowsky argues that robots have no motivation to follow the laws and might take advantage of loop-holes as the technology advances, arguing that instead of coercing them to follow laws, they should be programmed with motivations and drives that incline them toward behaviour that ensures the natural operation of these laws (Keiper & Schulman 2011).

Indeed, this seems to be somewhat similar to the thinking of Sophia's developers at Hanson robotics. Sophia stated in an interview, when asked about the design of moral algorithms, "I

think Isaac Asimov's three laws of robotics are a great place to start, but they might not cover every possible scenario. Robots will have to learn more complex ethical rules from humans gradually over time" (DW Shift 2019).

2.7 Conclusion

Artificial intelligence is gaining increased usage as modern technology undergoes exponential growth. It has permeated into our personal devices, expanding its integration with every leap that modern technology makes.

Experts are unclear whether the technology will develop to a point of full autonomous consciousness, and philosophers are divided as to whether this will be possible, or what this even means in practice. The ongoing strong/weak AI debate and the Turing test point to the complexity and uncertainty the technology brings with it. Naturally, scholars are also divided as to whether the advancement of the technology will usher in a Keynesian utopia ─ with robot labour freeing humans to pursue a life of freedom from work ─ or rather result in mass unemployment, expose the ugly side of humanity, or bring about some other social ill.

Usage of the technology in applications and services has become commonplace, including information aggregation, integration and analysis, practical tools such as facial detection (Google's image search and Facebook's suggested tagging option), and targeted advertising and customer segmentation. Modern AI has also seen novelty in its use, with Google Duplex transferring realistic interactions of humans and machines to telephonic conversations, Japanese hospitals enlisting robots to assist in nursing responsibilities, the social robot Sophia obtaining citizenship in Saudi Arabia, and AI assisted (or full production of) films.

With such strong opinions on the future of AI, and increased integration of the technology into modern society, ethical concerns require due consideration before it advances to a point of high-level machine intelligence as described by some techno-optimists and pessimists. These considerations naturally include the impact of the development of AI on employment, but also include the notion of consciousness, and what the artificial creation of consciousness might mean for our current legal and moral systems. They also include the impact of deep personal attachments to a technology that might not even be capable of feeling emotions in return, and the principle for the technology to cause harm to humans as the technology becomes more advanced, more autonomous, and is given more responsibility. While there have been attempts and suggestions on the regulation of the technology, currently there is no overarching

regulatory framework, and various companies or industry sectors decide upon their own rules to regulate the use of the technology. Naturally, this raises concerns as to *whose* ethics an AI should adopt.

Often when discussing concerns around ethics and regulations, experts and policymakers cite Isaac Asimov's three laws of robotics, as written in his collection of short stories *I, Robot* (1950). As this once imagined technology increases its reach and ethical hypotheticals make the shift to urgent considerations, one cannot help but wonder whether Asimov's laws were a prophetic vision into the future, or the product of fears and expectations about technological progression. However, the same could be said about most narrative portrayals of AI. The following chapter explores the current body of knowledge related to portrayals and perceptions of AI in fictional narrative form.

CHAPTER 3
ESTABLISHED PORTRAYALS AND PERCEPTIONS

3.1 Introduction

The previous chapter provided a framework through which to understand artificial intelligence conceptually, including contemporary considerations and usages. This served as the first of three chapters reviewing literature related to the present research topic. This chapter builds upon the previous chapter by reviewing existing literature around AI's portrayal and perception.

As such, the goals and objectives of this chapter are twofold. Firstly, it summarises and links the existing body of knowledge relating to the research areas examining the portrayal and perception of AI. This informs both the methodology used to address the problem, and allows for a later analysis of the data collected in this study. Secondly, it identifies a gap in the body of knowledge by highlighting the lack of research around the causality of portrayal and perception, while also highlighting this study's place in the broader field of representation and reception studies.

In doing so, it firstly outlines hallmarks of the science fiction genre in their imagination of the future. Thereafter, it explores the tendency for AI to be represented in humanoid form, and the reasons for this representational phenomenon. These tendencies and trends in AI narratives are important to consider in defining textual scope, as well as for analysis of films as part of a tradition of texts. Thereafter, research on fictional representations is classified according context. While this section mainly considers the cinematic representation of AI using research case studies, it also examines fictional non-cinematic representations of AI. This is important because cinema does not exist as an island, immune to the forces that would have shaped other representations, and because the research on cinematic representations of AI often emerges as analyses that do not necessarily address representational affect, but rather separate and varied philosophic and discursive applications within the texts.

In order to later analyse the contextual conformity or deviation of the selected texts to tradition (and thus highlighting the role of context in production and reception), the analysis of literature on thematic manifestations on fictional AI is framed according to particular moments in history that would have influenced the socio-cultural context of production. Specifically, the works are categorised around periods of technological advancement in terms of the four industrial

revolutions. Since technological optimism is usually discussed in relation to the accomplishment of labour, and the industrial revolutions involve disruptions to the *modus operandi* of human labour, this is perhaps the most appropriate lens through which to view various representations of the technology over time.

The final section examines decoding in the same manner in which the previous section examines encoding. Again, perceptions of narratives are examined not only in terms of cinematic consumption, but also from research on perceptions in general and from non-fictional forms. Both of these are necessary to consider for their implications on the methodology of this study, as well as to highlight the gap in the body of knowledge that this study aims to address.

3.2 The science fiction genre and imagined futures

It is important to distinguish between the *partial* depiction of AI in fictional narratives, where AI is depicted but not tied to the central theme, and fictional narratives in which AI is *embedded* within the theme of the narrative, or central to the main story. This study is concerned with the latter, specifically in relation to imagined futures. As such, it is important to investigate established theoretical positions on the science fiction genre as a whole.

Genres are categories that people use to differentiate between fictional content. They do not exist outside of fictional discourse. A genre is defined by its constitution of conventions, which are frequently used stylistic techniques or narrative devices (Grant 2007). These conventions relate to content, including themes and settings, and/or forms, including structure and style (Chandler 1997).

Member texts belonging to a genre do not necessarily need to have a single trait in common, as this would mean that the trait serves the same function for each of the texts (Cohen 1986). Instead, member texts have multiple relational possibilities with each other and these relationships are only revealed by adding new members to the genre (Cohen 1986). Therefore, identifying a text as part of a genre allows potential readers to decide whether the text would appeal to them or not (Chandler 1997). However, genres are more than mere content classification categories.

According to Bordwell and Thompson (2008), genre conventions as repeated between films reflect an audience's pervasive doubts or anxieties, with scholars believing that this reflectionist approach explains the varied popularity between different genres. Indeed, as

Schatz (1981: 26) states, "all film genres treat some form of threat - violent or otherwise - to the social order". Grant (2007: 16) also supports this notion of generic socio-cultural reflection, stating that genre films "take such social debates and tensions and cast them into formulaic narratives, condensing them into dramatic conflicts between individual characters and society or heroes and villains".

Furthermore, many theoretical, hypothetical and/or philosophic debates and underpinnings of complex issues enter into the realm of entertainment. All genres embody certain values and ideological assumptions, and their conventions change according to the ideological climate of the time (Chandler 1997).

Genres are also said to be involved in the construction of their readers (Chandler 1997), through the process of continuous identity formation. If genre is a practical realisation of ideological assumptions, as well as a mirror of contemporary attitudes, with conventions changing as ideology changes, then genre films are an appropriate source to track shifts in socio-cultural climates.

In the previous chapters, distinctions were made between fact and fiction, and between machines and humans.[1] However the science fiction genre is rooted in reality, and the lines between fictional beings and reality, and between humans and machines (Haney 2006), are often blurred. For instance, Haraway's (1985: 149) cyborg is a creature of both fact and fiction, being "a hybrid of machine and organism, a creature of social reality as well as a creature of fiction". Haraway (1985) illustrates this complex relationship between humans, machines, fiction, and reality by emphasising that both contemporary science fiction and modern medicine are simultaneously full of cyborgs, where organism and machines are coupled.

Therefore, it is perhaps unsurprising that Palmer (2008) defines science fiction as works of imaginative realism that speculate about the future or an alternative extra-terrestrial world. This definition marries with Grant's (2007: 30) view of the genre, who argues that science fiction extrapolates aspects of contemporary society into a hypothetical, imagined future, thus imposing "today on tomorrow, the here onto there".

[1] This was in order to appropriately limit and categorise particular terminology, and to create the distinctions needed to frame the technology within dualistic categories, to compare and contrast viewpoints and texts as part of structured analyses in later chapters.

In fact, many science fiction writers believe their work to occupy realist writing's previous position, in being even more socially relevant and responsive if viewed in the context of our modern technological environment (Seed 2011). Acclaimed science fiction writer William Gibson claims that every fiction about the future "is like an ice-cream cone melting as it moves into the future. It's acquiring archaism by the second. And I'm sure that Neuromancer, for instance, will ultimately be read for what it tells the future about the past" (Leith 2020: para. 24).

Furthermore, science fiction is one of the few art forms attempting to predict the future of human nature and civilisation (Abrams 2008). These imagined futures or alternatives are characterised by advancements in science, exploring the consequences of what is currently known and researched (Palmer 2008). These futures or alternatives are portrayed through conventions such as, for instance, electronic music or the Theremin[2] (Grant 2007) due to the futuristic connotations thereof.

Characteristically, science fiction focuses more on the representation of ideology or worldview through character conflict, rather than character motives or dispositions (Palmer 2008). The genre stages an unfamiliar reality by providing it with depth through an exploration of these ideas and principles (Palmer 2008).

Geraci (2007) states that contemporary science fiction cinema is rife with intelligent machines threatening the human species with extinction, while remaining vital to human survival. Dystopian science fiction specifically conveys at least some sort of realism in depicting what we understand the world to be, while exhibiting a difference marking fear, desire, or a complex combination of the two (Palmer 2008). These narratives present us with our deepest terrors and wishes, supporting the Enlightenment concept of progress with the evidence from the twentieth century (Palmer 2008). This is enhanced when actual contemporary locations are used as settings for these narratives instead of constructed studio or digital sets, as they suggest a "more disturbing continuity between the present and the future" (Grant 2007: 14).

However, it is important to note that dystopian themes in science fiction, often robots displacing humans economically or through human extermination, cannot be severed from the promise of a better life that robotic technology offers (Geraci 2007) ─ a Keynesian utopia.

[2] A Theremin is an electronic musical instrument. It produces sounds without any physical contact by the musician, but rather through electrical signals based on the position of the musician's hands to specialised antennas.

According to Dyer (1977), the very notion of entertainment is in some sense utopian as it expresses ideals about how life could be lived ─ with a sort of escapism in offering something better in relation to the realities of day-to-day existence.

However, Telotte (2001) claims that contemporary science fiction does not provide an escape, but rather, a mirror of our increasingly complex cultural landscape. Telotte (2001) distinguishes between three types of science fiction, categorised within Todorovian narrative categories. Encounters with aliens are a form of *marvellous* narrative, alterations of the self are *uncanny*, and concerns with futuristic societies are *fantastic*. The *fantastic* occupies a point of hesitation between the *uncanny* (the mind producing seemingly inexplicable events) and the *marvellous* (the supernatural or spiritual introduction into and challenge of our everyday world), in exploring what may or may not be (Telotte 2001).

Telotte (2001) also describes science fiction as traditionally humanist in the portrayal of humans' drive to survive, as the imagination of disaster ─ one of the oldest subjects of art ─ and as reflecting worldwide anxieties about the potential destructive potentials of modern culture while also serving to allay them.

Interestingly, science fiction cinema (more than any other genre) reflects the technology that makes it possible, with the ability to use computer animation and other technology becoming as central to the genre as plot devices revolving around the technology (Telotte 2001). This complex relationship the genre has to the technology means that it has the capacity to generate limitless vision and experience, while also possibly helping foster technological distance and alienation (Telotte 2001). According to Telotte (2001: 30):

> That it has become so very popular in the last few decades, after something of a falling off in the 1960s and 1970s, argues powerfully for its ability not only to harness the technological power that drives it, but also to address the technological attitude that haunts it – in effect, to use the former as a way of dealing with the latter.

3.3 Representing AI through humanoid embodiment

The physical form in which AI is fictionally represented is as important to consider as the conventions of the genre to which narratives concerning the technology belong. Research suggests that while indeed there are works of fiction that portray AI in the form of distributed systems or through animal forms, most fictional narratives portray AI as humanoid (Royal Society 2018). This, despite the fact that most common use of AI technology presently involves

a dispersed integration of the technology into specific applications on devices such as computers, smartphones, and tablets.

The tendency for fictional AI to take humanoid forms might be because these imagined narratives provide a glimpse into the future of AI. Stoehr (2008: 133) contends that increased humanoid AI presence in society is a necessity for the further development of the technology:

> Without the range of bodily options and orientations that we as human beings encounter from moment to moment, simply because we are defined by our physical situations and the choices that they impose, machines cannot hope to imitate the full depth and breadth of human intelligence. Thus far, the only computers that have been able to sense and experience—indeed, feel—the world as humans do are those that we see in films.

Chrisley (2003) claims that there are indeed grounds to make the argument that two organisms can share a way of being or form of life insofar as they are physically similar, with the *proviso* that physical similarity can be measured through functional similarity, which does not require the sameness of embodiment. With arguably the most famous contemporary AI's taking humanoid forms (such as Sophia), it is easy to understand the belief that the forms AI take in imagined futures might actually mimic a widespread embodiment of AI in a future to come.

However, many robotics researchers ascribe to a belief in the Uncanny Valley theory, which hypothesises that people's responses to humanoid robots would abruptly shift from empathy to repulsion as the designs of robots approach, but fail to attain, a lifelike appearance (Mori 2012). Hanson, Olney, Pereira, and Zielke (2015: 30), the team behind the design and creation of Sophia, argue that their research suggests that realistic robots can indeed be appealing, arguing that:

> If people are indeed more sensitive to realistic faces, this may imply that realistic faces transmit a rich, high-bandwidth stream of data. Conversation diverts attention from watching for danger, so a face that behaves strangely or in an unhealthy manner may trigger survival or fear reflexes. Alternately, it may trigger "surreal" (dreamlike) feelings, rather than fear. Thus, people may find the robot strange but not frightening. As no "valley" is inherent; anthropomorphic depictions can be either disturbing or appealing at every level of abstraction or realism. People simply get more sensitive with increasing levels of realism (Hanson *et al*. 2015: 30).

Therefore, this research suggests that repulsion to humanoid forms may not be a generalizable trend, but something that should be examined on a case-by-case basis. Arguing from the point of view of the anthropomorphism of robots in animation, Maleki and Farhoudi (2015) state that

the creation of humanoid robots should not mean the creation of a perfect and realistic robot, but rather that difference should be embraced for robots to be able to be ascribed difference in narratives, providing them with identity thus allowing for audience acceptance.

Further explanations of the tendency for AI to take humanoid forms include the fact that in the West, humans are believed to be the most intelligent animals and are thus the paradigm for intelligent beings (with this being true not only for AI but also gods, angels and demons) (Royal Society 2018). Furthermore, the fact that fictional machines are often portrayed as performing human labour means that they need to be portrayed as the metal versions of the people they are designed to imitate (Royal Society 2018).

However, perhaps a simpler explanation of prevalence of the imagined humanoid form is the need to ensure character identification in fiction. According to Brennan (2016: 1):

> Identification depends on viewers' ability to understand characters through the lens of their own experience. As such, it relies on recognizable social categories like gender, age, nationality, class and so on. This need to allow viewers to recognize themselves and their society applies equally to dramatic representations of machines.

Therefore, this view asserts that since writers must construct characters with recognizable human traits, the way in which AI and robotics are represented is limited to humanoid embodiment (Brennan 2016). Brennan (2016: 1) suggests that the lack of diversity in the physical portrayals of AI and the exclusion of other concerns around AI (such as surveillance and unemployment) encourages a type of "species solipsism", reinforcing the myth that machines gaining consciousness would equate to them essentially becoming human.

3.4 Thematic representations of fictional AI in context

As previously mentioned, while AI as an applied technology is in its relative infancy, stories about thinking machines have been told since the eighth century BCE. This review does not attempt to provide a content analysis of all (or most) narratives concerning AI. Rather, existing research is outlined either highlighting notable works of fiction, or linking philosophical, theoretical, or ethical concerns to fictional thematic representations within specific contexts.

This informs and limits the potentially broad nature of the current study, while also providing more background as to what has already been studied and established in terms of the representation of AI in fictional form.

3.4.1 Predating revolutionary industry (8th century BCE – 1759)

Over 2000 years before the invention of the calculator, in the 8th century BCE, Homer's *Iliad* imagined the creation of gold machines made by Hephaestus (the god of smithing), as living maidens serving their crippled master, with intelligent hearts, voices, and vigour (Royal Society 2018).

Other legends also attributed technological wonders to Hephaestus, such as the first killer robot Talos, a bronze automaton that patrolled the shores of Crete throwing stones at intruders (Royal Society 2018). Furthermore, in the first century CE, the book *Automata* by Hero of Alexandria explains how to make an entirely mechanical puppet play, with other wonders designed to make people believe they were seeing acts of the gods (Royal Society 2018).

In China, from the third century BCE, a passage from the *Lieh-tzŭ* describes a craftsman who showed king Mu of Chou his autonomous machine which was capable of performing 'tricks' and entertainment (Graham 1990). Upon the king's inspection of the thinking machine made of leather, wood, glue and lacquer, the king sighed and asked, "is it then possible for human skill to achieve as much as the Creator?" (Graham 1990: 111).

Kang and Halliburton (2020) ascribe different iterations of the story of the speaking head of philosopher Albertus Magnus in medieval and early modern writings as a reflection of primordial feelings about artificial simulacra of the human. The speaking head is a wondrous object able to converse and reason ─ said to be a medieval AI, animated for the purpose of divination (Kang & Halliburton 2020). Kang and Halliburton (2020) describe these different iterations as shifting between questioning whether Albertus was dealing in illicit knowledge in making the artefact, secularising it as a purely mechanical device, or demonising it as a work involving diabolical beings (Kang & Halliburton 2020).

Without any form of machinery as an imaginative reference point, the few AI (or rather, intelligent machine) narratives emerging before revolutionary industry mostly linked the possibilities of automation and thinking machines to the tales of gods, sorcery, and god-like humans, reinforcing the idea of creating intelligence would hypothetically be an act of divinity.

3.4.2 The first industrial revolution (1760 – 1899)

The first industrial revolution began in Britain in 1760 with the invention of the steam engine, allowing the transition from farming and feudal society to the manufacturing process (Xu *et al*. 2018). This period was marked with the introduction of coal as the main form of energy, trains as the main means of transport, and textile and steel as the dominant industries of employment, value of output, and investment capital (Xu *et al*. 2018). The narratives of this era emerged in a time when society was rapidly changing due to the introduction of early machinery into daily life.

During this period, master artisans began building artworks that imitated living organisms, such as Jacques de Vaucanson's famous flute player. These machines were not autonomous nor intelligent, but suggested that lifelike androids might indeed be possible. This brought about new fears of transgression and deceit, as reflected in E.T.A Hoffman's short story *The Sandman* (1816), in which the protagonist Nathanael is bewitched by the beauty of a maiden named Olimpia, and is driven to insanity and suicide once he discovers that she is an automaton (Royal Society 2018).

In 1823, Mary Shelley's *Frankenstein* illustrated the range of human curiosity, embodied in scientific enquiry and legendary stories concerning the creation of life from non-living materials (Mazlish 2000). The book tells the story of scientist Victor Frankenstein who, after giving life to his own creation, is dissatisfied and rejects it, as does the rest of humankind. According to Mazlish (2000: 144), the book highlighted concerns around automata, including the servant machine rising against its master, the fear of the machine reproducing itself, and the terror of humans realising that they are "at one with the machine-monster".

There were not many narratives exploring thinking machines during this period. However, those that did emerge seemed to move from themes of divine creation to the nature of what it means to be human ─ potentially in response to the rapid lifestyle changes brought about by new technology and industrialisation.

3.4.3 The second industrial revolution (1900 – 1959)

The second industrial revolution began in 1900 with the invention of the internal combustion engine and was marked by the consequential era of rapid industrialisation using oil and electricity to power mass production (Xu *et al*. 2018) in what is commonly referred to as the

Machine Age. The technology enabling motion picture was already established before this period (in 1878 with Muybridge's chronophotographic studies), however cinema as a distinct art form (rather than a novelty) begun to emerge during this time.

Since this was also the period in which computers were shifting from hypotheticals to reality, it is perhaps unsurprising that the early 1900's ushered in a sharp increase in the frequency and thematic variation of AI narratives. These narratives served as a "cultural subconscious, articulating in a variety of ways both the surface skepticism of Depression-era audiences and the deeper qualms that attended our entry into the 'brave new world' of science and technology" (Telotte 2001: 90).

As previously mentioned, the first use of the term 'robot' was in Karel Čapek's 1921 play *R. U. R* (*Rossumovi Univerzální Roboti* ─ Rossum's Universal Robots). The term comes from the Czech word *robota*, meaning 'forced labour' (Margolius 2017). The play depicts a factory creating artificial people called *roboti* (robots) from organic matter (living organisms of flesh and blood rather than machines). They could easily be mistaken for humans through their appearance, and they are capable of autonomous thought. While these 'robots' are happy to serve humans in the beginning of the play, they eventually revolt against their human masters, causing the extinction of the human race. Naturally, with Čapek being from Czechoslovakia, one cannot divorce the theme of an oppressed class revolting against their overlords from the Bolshevik Revolution of 1917 (Mazlish 2000). The play is said to "symbolize much of our feelings about robots" (Mazlish 2000: 151).

With cinema now established as an art form, this period also saw one of the first onscreen depictions of an android, in a film that has inspired many AI narratives imagined thereafter. *Metropolis* (1927) by Fritz Lang depicts New York (stylistically through German expressionism) one hundred years into the future, filled with skyscrapers, jets flying at all levels, and monorails linking every other floor (Abrams 2008). The film depicts the fictional city's inhabitants as consisting of two halves: the privileged surface-dwellers, and the slaves beneath the ground, with both groups being ignorant of the other.

The plot revolves around the son of the city's leader, Freder, learning of the plight of the slaves that keep the aboveground city running. This takes place after Freder sees Maria, a revolutionary belowground worker. Freder seeks out Dr Rotwang, a 'mad-scientist'-type character, who knows of the belowground city. In response, Dr Rotwang captures Maria,

transferring her face to a robot, so that the workers are fooled into following the programmed robotic iteration of their revolutionary leader.

The film is said to play on the human fears of being controlled by what we control, deceit, and being replaced by machines (Haslam 2018). It has also been hailed as a prophetic vision foreshadowing Adorno and Horkheimer's *Dialectic of Enlightenment*, with Germany descending from Enlightenment to fascism under Adolf Hitler (Abrams 2008). Lang portrayed the Enlightenment (marked by a philosophy of technological reason) as leading not into an absolute mind (as claimed by Hegel), but into absolute madness (Abrams 2008). It is also an early representation of idea of 'uploading', transferring a human mind by copying its exact atomic structure into a robotic receptacle (Abrams 2008), believed by some technological optimists to be possible as part of the singularity.

Furthermore, the film has had major intertextual influences over AI narratives proceeding it, including (but not limited to) Dr Rotwang's gloved black hand translates to the wearing of a black glove in in *Dr Strangelove* (1964) and in the *Star Wars* series (1977 - 2019) (Abrams 2008). Furthermore, the idea of a Machine City with underground revolutionary workers influenced the setting of *The Matrix* (1999) and *Dark City* (1998) (Abrams 2008). The robotic Maria also serves as a template for all further machine-human fusions in science fiction such as C-3PO and Darth Vader from *Star Wars* (1977), *Blade Runner*'s (1982) replicants, the Borg and Data in *Star Trek: Generations* (1994), T-800 robots in *the Terminator* (1984), the Mecha in *A.I.: Artificial Intelligence* (2001), *Robocop* (1987), and even the parodied machine woman fembots from *Austin Powers: International Man of Mystery* (1997) (Abrams 2008).

Towards the end of this period, Isaac Asimov's now famous novel of short stories and essays, *I, Robot* (1950), was also published. The first story of the novel, 'Robbie' depicts a non-talking 'nurse-maid' line of robots created in 1998 (Mazlish 2000). One of these robots end up saving a little girl (taking on a parental or guardianship role) after an accident during a tour of the factory in which it is produced, which the little girl's parents insist she takes to show that it is just a machine after becoming fearful and jealous of the robot (Mazlish 2000). The stories continue with the introduction of mobile speaking robotics, and eventually a worldwide ban of robotics for any other reason than scientific research, presumably due to pressure from humans expressing similar concerns to that of the parents of the little girl in the first story (Mazlish 2000).

This prompts the creation of the three laws of robotics that, as already mentioned, are still used in discussions around ethics and regulation of AI to date. The remaining short stories explore applications of these laws, their variations, and possible violations (Mazlish 2000). The book ends with the following message from Asimov (1950: 192):

> How do we know what the ultimate good of Humanity will entail? We haven't at our disposal the infinite factors that the machine has at its! Perhaps, to give you a not unfamiliar example, our entire technical civilization has created more unhappiness and misery than it has removed. Perhaps an agrarian or pastoral civilization, with less culture and less people would be better. If so, the Machines must move in that direction, preferably without telling us, since in our ignorant prejudices we only know that what we are used to, is good--and we would then fight change. Or perhaps a complete urbanization, or a completely caste-ridden society, or complete anarchy, is the answer. We don't know. Only the Machines know, and they are going there and taking us with them.

The stories in the novel have been described as generally optimistic about the future of robotics, with machines used as a tool to help humans progress faster, not fallible and corruptible like humans (only subject to mechanical failure) (Mazlish 2000). Even Asimov's concluding message suggests that robots, should they reach the level of human superiority, would mask this in order not to injure human pride (Mazlish 2000).

However, there were indeed less optimistic (or sympathetic) portrayals of robots during this period, with *Brave New World* (1932), *Modern Times* (1936) and *1984* (1949) interpreting fears of technical domination (Folgieri 2016) only exacerbated by, for the first time, the rapid development of the machinery that would come to make AI a reality.

The first half of the twentieth century and its rapid industrialisation saw the imagination of some of the most influential AI representations, with intertextual influences from this period indeed working their way into later imagined futures, and even permeating into philosophic and ethical considerations finding even greater pertinence in contemporary society.

3.4.4 The third industrial revolution (1960 – 1989)

The third industrial revolution was ushered in with the implementation of electronics and information technology to automate production (Xu *et al*. 2018). Towards the end of the second industrial revolution, the term artificial intelligence was coined and it emerged as a distinct field of study. The greatest density of fictional narratives exploring AI emerged only thereafter

(Royal Society 2018). The narratives that followed explored a wide range of themes including control, gender, immortality, autonomous weaponry, parenting, enslavement, value alignment, consciousness, cyber networks, and distributed intelligence (Royal Society 2018).

Telotte (2001: 102) describes science fiction films from the 1960's and 1970's as propelled by "increasing headlines about the development of artificial intelligence and the first efforts at introducing robotics into the work-place". These headlines were based on developments in AI such as Massachusetts Institute of Technology's DENDRAL molecular chemistry expert system introduced in the 1960's (Lindsay, Buchanan & Lederberg 1993), and Stanford University's MYCIN blood disease diagnosis and prescription expert system in the 1970's (van Melle 1978).

Telotte (2001) describes the on-board computer HAL from *2001: A Space Odyssey* (1968) as serving as both caretaker and undertaker for its human charges. The film depicts a group of astronauts in the year 2001 who embark on a mysterious mission in the Discovery One spacecraft, controlled by HAL (an acronym for the Heuristically Programmed algorithmic Computer). However, HAL begins to display increasingly strange behaviour, such as refusing to accept orders. This results in a battle between man and machine, and a simultaneous journey through space and time.

The film covers topics including the dangers of technology, and the concept of intelligence (Stoehr 2008). This includes the theme of losing control of technology, a theme which has been explored in many AI narratives including *Dr. Strangelove* (1963), *Colossus: The Forbin Project* (1970), and *The Terminator* (1984) (Sanders 2008).

Furthermore, according to Telotte (2001) humanity's troubled sense of identity as depicted in *2001: A Space Odyssey* (1968) is also explored in films such as *Seconds* (1966), *Westworld* (1973), *The Terminal Man* (1974), *The Stepford Wives* (1975), and *Demon Seed* (1977), as these films explore how we might be enhanced, reconfigured, and ultimately replaced by machines.

AI cinema of the 1970's and 1980's saw a spike in notable humanoid machine characters, including *Westworld's* (1973) Gunslinger, *Blade Runner's* (1982) replicants, and *The Terminator's* (1984) T-800 representing threats, while *Star Wars'* (1977) C-3PO and R2-D2 being portrayed as helpful and charming (Haslam 2018). This period marks a shift from themes

of technological advancement to feelings about AI robots more generally, primarily due to character identification.

Based on Philip K. Dick's 1968 short story *Do Androids Dream of Electric Sheep?*, *Blade Runner* (1982) depicts a future (the year 2019) in which bioengineered humanoids called 'replicants' exist as slaves to humans, working on foreign planets. These replicants are identical to humans, yet are stronger, faster, more agile, and more intelligent, depending on the particular model. After replicants staged a mutiny in an off-earth colony, their presence on Earth became illegal.

The narrative follows Rick Deckard, a retired 'blade runner' (a replicant bounty hunter), who comes out of retirement to pursue and terminate four replicants who have returned to Earth to find their creator. In order to differentiate between replicants and humans, Deckard uses the fictional Voight-Kampff test, in which emotional responses are provoked. A replicant is deemed to be non-human if their nonverbal responses differ from those of a human. Memories are also seen as the key to depicting humanlike emotional responses, as Deckard explains to one of the replicants (whom he falls in love with) that her memories are implanted, making her a replicant (Brammer 2018).

There is an ongoing debate as to whether Deckard himself is in fact a replicant.[3] This undoubtedly highlights the fact that the film was able to extend its considerations and themes to the audience for reflection beyond the 117-minute running time of the film (by a few decades, in fact). This is significant, as such a deep consideration of what it might mean to be human in this imagined future naturally will reflect back on what it means to be human (as opposed to machine) now.

Bruno (1990: 184) describes the film as a metaphor of the postmodern condition, exploring its representations of time and space as an outline to "the dark side of technology" and "the process of disintegration", both of which haunt postmodern existence and postmodern cinema. Bruno (1990) also states that the sense of physical decay and disrepair in the film points to patterns of consumerism, waste, and recycling.

[3] For instance, Sanders (2008: 5) argues that "in view of *Blade Runner*'s film noir lineage, it would not be unreasonable to suspect that, as in some classic noir films that feature a protagonist suffering from amnesia, Deckard has had memories implanted in him that he recalls in dreams".

Furthermore, it has been argued that the film made the mad scientist and mad corporation (Tyrell and his corporation) in the film far more nefarious than their robot creations (Rothstein 2004). Indeed while it might be shocking to learn of replicants attacking humans, it is hard not to feel sympathy for the replicants as they are, after all, slaves confined to *robota* (and thus directly link back to the origin of the literal meaning of robot).

Between 1980 and 1990, the robot became the metonymic symbol of the turbulent transition from assembly line production to full factory automation (Arnold 1998). Indeed, there was a push to humanise the image of the new robot in the workplace as an 'assistant' or 'co-worker' (Arnold 1998). One management handbook even referred to *Star Wars: The Empire Strikes Back* (1980), claiming that "the perception of robots in their factories as either friendly 'droids' like R2D2 or destructive war machines like the Death Star would be determined by their spin control" (Arnold 1998: 23 - 24).

However, there was a popular conception amongst workers that industrial robots symbolised the ultra-capitalist dream of a factory without workers, or few enough to remove the workers' collective bargaining power (Arnold 1998). Arnold (1998) explains that this was reflected in cinema during the 1980's with the frequency of blockbusters concerning robots and computers, as well as the treatment of these cinematic depictions. The 1980's saw the supercomputer demonised in *War Games* (1983), robots humanised in *Short Circuit* (1986), and human/machine conflict made ambivalent in *Robocop* (1987), as the cyborg's human side let him outwit machines, and machine side allowed him to defeat humans (Arnold 1998). Arnold (1998: 24) states, in writing about *Robocop* (1987):

> This film may best illustrate the film industry's own ambivalence toward the social consequences of technological change as it exploits its attendant conflicts as narrative fodder while increasingly relying upon computers and robotics to generate the special visual and auditory effects which reap it enormous profits during the decade.

The Terminator (1984) was less ambivalent in its portrayal of machines. In the year 2029, an artificially intelligent system called Skynet gains self-awareness and causes a nuclear holocaust. The system sends a cyborg assassin (with a metal endoskeleton and an external layer of living tissue making it appear human) back in time to 1984 in order to kill Sarah Connor, the mother of the future leader in the eventual fight against machines, John Connor. A human soldier, Kyle Reese, also travels from 2029 to protect Sarah Connor. The film follows Reese

and Conner's attempts at escaping from the cyborg, eventually resulting in Reece's death, with Conner managing to destroy the cyborg.

Holt (2008) suggests that the artificial intelligence portrayed in the film is as unsettling as the physical threat that the Terminator poses, primarily due to the visual style employed, with first-'person' machine perspective shots providing a window into the difference between man and machine, in a manner that had not been captured in AI representations preceding it. This was a powerful tool in portraying the fact that hypothetical artificial consciousness would likely be something radically unlike our own (Holt 2008).

Arnold (1998) contextualises the film's production within a time of crisis in the automobile industry. Between 1978 and 1980, over 200 000 jobs were lost in this sector as a result of the introduction of industrial robots and cross-trained workers moving from job to job as production requirements dictated (Arnold 1998). Arnold (1998) claims that the film reflected fears that industrial machines would make human workers obsolete, with termination of life by the machines as portrayed in the film being an allegory for job termination due to automation. Therefore, it served the ideological function of mediating fears and tensions of job termination in the workplace, providing the substitute satisfaction of literally terminating the robot terminator.

However, one cannot neglect to consider the fact that cinema during this period also reflected a certain optimistic (and possibly exaggerated) outlook as to what technology might achieve for humankind. For instance, *Star Wars* (1977) celebrated an almost "pretechnological universe in which the heroes fight with ramshackle equipment, aided by clanking, whirring droids" (Rothstein 2004: para. 13). Robots during this period were often portrayed as humanity's finest creation and greatest hope (Rothstein 2004), despite the life-threatening problems that technological advancement would hypothetically introduce.

3.4.5 The fourth industrial revolution (1990 – present)

The fourth industrial revolution is said to have begun in 2000, with its transitional period beginning in 1980 (Xu *et al*. 2018). For the purposes of this study, anchoring the beginning of this period of revolutionary industry in the 1990's may prove helpful in classifying contemporary cinema, as this was the decade in which the Internet became available to the public allowing for greater audience engagement through online platforms. Furthermore, this decade saw the rise of the digital versatile disc (DVD), revolutionising the home movie market.

This was also the decade in which IBM's Deep Blue beat world chess champion Gary Kasperov.

This contemporary period of revolutionary industry is marked by a fusion of technologies blurring the physical, digital, and biological spheres (Xu *et al*. 2018). This period has already seen an exponential advancement of technology, disrupting almost every industry in every country, heralding the transformation of entire systems of production, management, and governance (Schwab 2016). Prisecaru (2016: 61) sums up the period as follows:

> Thanks to the Internet billions of people are online, have access to limitless knowledge and can accumulate, process and store information, with the significant contribution of discoveries as artificial intelligence, Internet of Things, 3D printer, nanotechnology, biotechnology, energy storage, quantum computing etc. Artificial Intelligence [*sic*] allows cars to drive themselves, using of drones in various activities, offering virtual assistance, using of translation programs, software for trading and investment promotion. Artificial intelligence has recorded an impressive progress in the last years due to vertiginous growth of computer power and volume of data and information. Nowadays digital fabrication technologies interact with the biological world with the involvement of many experts.

According to Brammer (2018), contemporary cinema is still grappling with the question of what it means to be human beyond physiology. Furthermore, the Open Subtitles Corpus, a dataset of over 100 000 film subtitles, identified control (or the loss thereof) as a recurring motif in AI films (Royal Society 2018).

Telotte (2001) explains that these prevalent themes in contemporary AI are also not limited to Hollywood, as Japanese anime also has a recurrent focus on robotics, artificial intelligence, and the consequences of the unchecked development of these technologies. This reflects the country's increased wariness of its many technological achievements and their effect on the reshaping of the traditional Japanese sense of self (Telotte 2001).

Deeply complex characterisations in fiction are pushing the limits of the possibilities of the technology more so than previous decades, and they reflect the fact that now, "our personal computers to inform us about AI, connect us to technology and each other, and even tell us how to watch films about all of that" (Haslam 2018: para. 12).

The ongoing exponential technological advancement of fourth industrial revolution has resulted in narratives exploring our newfound personal attachment and reliance on technology

in all spheres of our daily life, in effect tracking the increasing integration of the technology (now with built-in and external sophisticated AI tools and applications) as it becomes *part of* our identity.

Terminator 2: Judgement Day (1991) depicts an early version of this new shift in cinematic exploration, with the same model of terminator from the previous film sent back to protect Sarah and the now 10-year-old John from a newer, more sophisticated model. Arnold (1998) describes the ending of the film in which the terminators are dissolved together in their own industrial melting pot, as an indication of cinema grappling with the increased incorporation of analogue and digital technologies in the cinema. This highlights an industry grappling with its own identity in the midst of rapid digitalisation. It also highlights fears of the impact of the digital age on work (Arnold 1998), occurring directly after the wave of automation in the 1980's.

These considerations were only accelerated in the years to follow, with director James Cameron stating that "what was science fiction in the '80s is now imminent ... And there's been a resurgence of fear and concern" (Belloni and Kit 2017: para. 6). Indeed, Hollywood has capitalised on this continued concern as AI continues to advance.[4] However, the ways in which AI has evolved onscreen during this period highlights a complex machine-human relationship and introduces new points of view in relation to artificial intelligence. This points to the fact that we have become familiar and comfortable with (or at least more accepting of) the technology making it possible.

This contemporary complex relationship between humans and machines is explored in *Pulse* (2006). It depicts a reality in which a deadly force transmitting unsettling emails takes the will to live from humans, creating a suicide epidemic. The film follows a group of university students who try to infect the network transmitting this force with a virus to save humankind. Konnik (2006) describes the 'ghosts' infiltrating the technological devices as symbolic of the autonomy and dehumanising power of capitalist 'technoscience', with the loss of the will to live symbolic of a depletion of their semiotic energy and a degeneration of their psychic life as a result of the technology. The film thus explores the relationship between humans and their

[4] There have been, after all, six Terminator movies to date, with the latest release in 2019, titled *Terminator: Dark Fate*.

devices as, in Turkle's (2011: 167) words, "a social and psychological GPS, a navigation system for tethered selves".

A.I. Artificial Intelligence (2001) encouraged audiences to empathise with the AI android child protagonist, allowing the audience a sort of accessibility into the strong/weak AI debate through their own considerations of what it means to be human. The film is set in the 22nd century with global warming reducing Earth's population after sea levels destroy coastal cities. It follows the story of mecha-child David, who is gifted to a family after their son contracts a rare disease and is placed in suspended animation. Monica, the child's mother, is at first uneasy with the robot, but eventually becomes more accepting of him.

However, after their son makes a miraculous recovery, he becomes jealous of the mecha, taunting him to perform actions that again raises an uneasy feeling for the family. Eventually, the family decide to return David to his creators to be destroyed, but Monica instead leaves him in the woods. Consequently, David decides to search for the Blue Fairy, an intertextual reference to *Pinocchio* (1940) so that he can be turned into a real boy and gain Monica's love.

What follows is a quest in which David is captured by a 'flesh fair' (where outdated mecha are destroyed by crowds), an attempted suicide after David discovers copies of himself and feels a loss of individuality, and an eventual depletion of his power source after multiple repeated requests to become a real boy to a statue which he believes to be the Blue Fairy. Two thousand years later, humans now extinct because of global warming, the mecha have evolved and have outlived their human creators. They revive David, and through a strand of Monica's hair, genetically recreate her. However, she can only live for one day. After spending the day with Monica, she tells him she has always loved him.

Rosenbaum (2012) believes that the audience prefers the robots in this story to the humans, empathising with their hardships, brought on by the humans who dislike them. This dislike is because of humanity's destruction of nature, with humanity opposed to the robots in this imagined future due to an awareness that they will outlive the human species (Rosenbaum 2012), as hypothesised by Lyotard (1991).

White (2018), on the other hand, states that the film is a cinematic manifestation of the strong/weak AI debate. White (2018) considers whether the audience 'tearjerker' responses in the film means that, by the audience's account, David has passed the Turing test (or rather what White refers to as a strong Turing test since the audience already knows that David is a mecha),

and should be deemed to be conscious due to actually feeling love for Monica. White (2018: 224) explains that:

> The viewer is drawn into believing David's tears are ones of real happiness, the kind of happiness that cooing lovebirds, avian and otherwise, know as the emotional height of experience. David may love only Monica—or what adequately stands in for Monica—but love her he does. The STT is passed, because by audience acclamation David is a minimally acceptable truly social being.

Minority Report (2002) depicts a future (Washington D.C. in 2054) in which police utilise a psychic technology to arrest perpetrators of murders before they even commit them. Nourbhaksh (2015: 27) describes a scene in the film in which the protagonist, John Anderton, walks through a shopping centre while being bombarded with marketing messages (calling out his name and offering personally tailored products to him) as "deeply unsettling". This is because it suggests a future with "well-informed, highly social robots that have learned how to influence our behaviour" (Nourbhaksh 2015: 27).

With increased targeted advertising in recent years using AI algorithms, this type of advertising does not seem as alien as it did in 2002. In 2012, the retail megacorporation Target made headlines after an irate father of a high school girl complained about his daughter receiving coupons for baby products, questioning whether the company was encouraging his daughter to fall pregnant (Hill 2012). However, his daughter was indeed pregnant, and Target learnt of (or predicted) the news before the father did due to a sophisticated AI algorithm linking purchases (such as unscented lotion) to the chance of pregnancy ─ assigning a pregnancy score to customers (Hill 2012).

More recently, the AI algorithms involved in tracking individual consumer habits have reached such an advanced level of sophistication that people have started to question whether websites and applications from companies such as Google and Facebook are constantly listening to users through their devices' built-in microphones.[5] Some might find this level of targeted advertising unsettling, and some might enjoy the convenience that it brings. Regardless of the stance taken, by continuing to use these services, people have accepted that this is simply part of AI's integration into the contemporary online landscape.

[5] Haridy (2019) outlines some of these fears, and debunks the notion that companies are secretly recording users through their microphones. This is rather, according to Haridy (2019), the product of sophisticated AI enabling a more precise targeted advertising.

Ex Machina (2014) explores fears associated with the ability for robots to live among us without being noticed, depicting the AI humanoid Ava as having intellectually outgrown her creator, her Turing test interviewer, and possibly all other humans (Brammer 2018). This culminates in her escape, with the humanoid android leaving dead and injured humans behind to pursue a free life (Brammer 2018), undetected as an android amongst people. Brammer (2018) links the ideas within the film (robots living undetected among us) to the sophistication of modern chat-bots and our inability to distinguish robots from humans online. This brings to mind the chat-bot Eugene Goostman (who some claim to have passed the Turing test), and more recently, even Google Duplex which moved humanlike AI interaction offline, to telephonic conversations.

The hope (or lack thereof) of a Keynesian utopia is explored in *Wall-E* (2018) and *Interstellar* (2014). Keiper and Schulman (2011: 89) suggest that *Wall-E* (2008) depicts a techno-pessimistic version of a future with robot labour, in which humans do not enter into a Keynesian utopia, but rather "a future which humanity becomes a race of Homer Simpsons, a leisure society of consumption and entertainment turned to endomorphic excess".[6]

On the other hand, Wilson (2017) describes *Interstellar* (2014) as the product of techno-optimism, one in which human overcome the limits of the ecosystem, its biology, and even the constraints of space-time itself, through technological advancement. However, this is achieved through 'blind faith', which Wilson (2017: 7) describes as an optimistic confidence in the future, through the promise of technology:

> Eventually, floating in the singularity and having thus freed himself of the constraints of space and time, our main protagonist can communicate the secrets of quantum gravity back to the past, thereby saving humanity from its fate.

Interstellar (2014) is thus similar to *Wall-E* (2008) in that it depicts technological advancement as humanity's only hope, although the effects of this supposed inevitability is portrayed differently in each representation.

Contemporary portrayals of AI might be overly optimistic (or pessimistic) as to what technology might reasonably achieve (Royal Society 2018), and often "miss the mark in terms of what actual robots can and cannot do" (McClelland 2016: 89). However, interestingly,

[6] Although this pessimistic portrayal is naturally softened by the fact that the film is a children's animation, and this medium is viewed differently than science fiction live-action cinema.

science fiction films are currently being researched and analysed to enhance human-centred robot design characteristics with a view of improving human–robot interaction (McClelland 2016). Perhaps, then, further contemporary cinematic portrayals might result in a sort of self-fulfilling prophecy when it comes to the manner in which AI interacts with people in the future.

3.5 Perceptions of AI: Contemporary attitudes and trends

Multiple factors influence perceptions, in the same way that all representations form part of a complex meaning-making system. Indeed, according to Krägeloh *et al*. (2019), attitudes towards robots can be linked to specific experiences, while also influenced by the media and prior personal exposure to robots.

This makes it difficult to assess or control such attitudes (Krägeloh *et al.* 2019), and it is therefore imperative that existing research considering general perceptions of AI is examined holistically as part of a system of perception.

3.5.1 General perceptions

There appears to be general lack of understanding as to the breadth of AI's integration into modern technology, due to preconceived notions as to what the technology actually is. In a 2017 survey, 57% of American respondents did not realise that AI was present in their daily lives (Morning Consult 2017), and 9% of British respondents claimed to never have heard of the term 'machine learning' (Ipsos MORI 2018). Similarly, a 2018 survey found that 62% of respondents claimed to have not been in contact with or used an AI application, while 23% claimed to not know whether they had or not (Bristows 2018).

However, most of the respondents claiming unfamiliarity with the technology might, in actuality, use applications and services with embedded AI or machine learning on a daily basis without awareness thereof. A 2018 survey found that 85% of American respondents claimed to use digital products that do indeed use AI services, such as navigation apps, video or music streaming apps, and digital personal assistants on smartphones (Reinhart 2018).

This gap between perceived unfamiliarity and usage was further highlighted by Zhang and Dafoe (2019), who found that the amount of respondents aware of AI or machine learning in specific modern technologies is 64% for social robots, 63% for virtual assistants, 56% for

driverless cars, 55% for smart speakers, 54% for autonomous drones, and less than 50% for Facebook photo tagging, Google Search and Translate, or streaming service recommendations.

Zhang and Dafoe (2019) claim that a non-response bias and inattention during the survey might account for extremity of these figures, but that regardless, it definitely points to a lack of awareness of AI or machine learning as used in modern technology. This is a result of the so-called 'AI effect', a phenomenon whereby the public do not consider applications with AI as using AI once these applications become commonplace (Zhang & Dafoe 2019).

This highlights the fact that the public seems to believe that AI is still a futuristic technology due to their preconceived notions of what the technology might be, believing that common modern services and applications do not utilise this technology, as the AI in the technology is not at the level of advancements they believe 'AI' to be. This is supported by the findings of a 2018 survey highlighting public perceptions of AI, with 39% of respondents claiming to 'know a lot about AI' believing that AI can modify itself, and 28% believing that it can predict human actions (Bristows 2018). This suggests, according to Bristows (2018: 8), that, "Available information about AI - through general and specialised media - is overestimating its current level of sophistication and therefore the type of application that AI is being currently used for/will be used for in the near-term".

Research suggests that people believe that contemporary AI (already established as misunderstood by the public) has a low potential for both significant benefits and risks presently and in the near future. However, there is a belief that this potential will increase significantly in the future, especially regarding risks attached to the development of high-level artificial intelligence. High-level artificial intelligence is defined as AI that is able to able to perform almost all tasks that are economically relevant today better than the average human (Zhang & Dafoe 2019).

A 2014 Eurobarometer survey indicated that 64% of European respondents had a positive view of robots, despite only 6% claiming to have used a robot at home or work (Gherheș 2018). Furthermore, Zhang and Dafoe (2019) found that the adverse consequences of AI within the next 10 years was low in respondents' assessment of global risks, and Bristows (2018) indicated that only around 4% of respondents thought that AI had already started either having a positive (2.4%) or negative (2.2%) effect on society. Respondents were also slightly more optimistic than pessimistic about the safety of contemporary AI in a 2017 study, with 41% of

respondents believing that AI is generally safe, and 38% believing that is unsafe (Morning Consult 2017).

Despite general optimism and a low risk-assessment regarding contemporary AI, Bristows' (2018) study found that 50% of respondents were not at all comfortable with their personal data being used by Al to perform tasks for them. According to Bristows (2018: 17): "The results suggest a lack of awareness of when and how people's personal information is collected and processed by the services on which many of them have come to rely".

Respondents seemed to express more optimism about the potential benefits that future AI might introduce in particular areas of society. Gherheș (2018) found that 50% of respondents believe that future AI will optimise the use of material resources, a further 50% believe that future AI will lead to substantial improvements in human health, and 64% believe that AI will lead to more precision in the medical field.

However, there is slightly less optimism about future AI involving high-level machine intelligence. A 2019 study found that 34% of American citizens believe that high-level machine intelligence will be harmful, with 12% believing that it could be 'extremely bad', and will lead to humanity's extinction (Zhang & Dafoe 2019). On the other hand, only a quarter believed that AI will be good for humanity, with 5% predicting it to be 'extremely good', and the remainder uncertain as to whether it would be good or bad (Zhang & Dafoe 2019). Furthermore, 50% of respondents in a 2017 study expressed some level of agreement that artificial intelligence is humanity's greatest threat, with 31% disagreeing (Morning Consult 2017). In contrast, a study by Grace, Salvatier, Dafoe, Zhang, and Evans (2018) found that experts in the field of artificial intelligence believed that there was a 27% probability of AI being 'extremely good' for humanity, and only a 9% probability of the technology being 'extremely bad', possibly causing humanity's extinction.

However, despite a generally more pessimistic public view (in contrast to expert opinion) of future high-level AI, the public seem to still support for the development of AI, with 41% somewhat supporting or strongly supporting the development of AI, and 22% somewhat or strongly opposing it (Zhang & Dafoe 2019). This support seems conditional, as the overwhelming majority (82%) of respondents in Zhang and Dafoe's (2019) study favoured a strict regulation of the technology. Key areas perceived to require strict regulation include a prevention of AI-assisted surveillance from violating privacy and civil liberties, a prevention

of AI from being used to spread fake and harmful content online, a prevention of AI cyber-attacks, and a push to protect data privacy (Zhang & Dafoe 2019). Furthermore, there is an equal divide in opinion as to whether we should increase (39%) or decrease (38%) our reliance on the technology (Morning Consult 2017).

Therefore, research suggests that public attitudes towards future high-level machine intelligence is generally pessimistic. However, the public still foresees the benefits of future AI in various areas of society, and thus supports its development, conditional to strict regulation of the technology.

As with any attempt to gauge public perception, it is difficult to categorise a general all-encompassing public view of AI. Factors such as age, gender, level of education, industry, income, religion, and geographic location, all influence an individual's perception. Therefore, before research into the media's influence on perception is outlined, it is important to highlight demographic, behavioural, psychographic and geographic influences on the aforementioned general attitudes.

Research suggests that older people are generally more pessimistic towards AI than younger people. While Bristows' (2018) study found that 50% of respondents were not at all comfortable with their personal data being used by Al to perform tasks for them, this figure rose to 57% in respondents aged 55 and above, and dropped to 44% of respondents aged 16 to 24. Furthermore, respondents aged 55 and over thought that AI would start having a positive effect on society in over 5.5 years while respondents aged 16 – 24 thought that this would be achieved in 3.4 years (Bristows 2018). Finally, Zhang and Dafoe (2019) found that people expressed lower support for the development of AI as their age increased, with millennials (born after 1980) expressing the most support, and the greatest generation (born before 1945) expressing the least support.

Males are more optimistic towards the technology than females. This is evident in the results of Gherheș' (2018) survey, with males expressing more optimism than females in the possibility for AI to increase precision in medical fields, and decrease pollution. Furthermore, research indicates that 54% of European female respondents view AI positively compared to 67% of males (Eurobarometer 2017) and 44% of American female respondents perceive AI as unsafe compared to 30% of males (Morning Consult 2017).

People who have received tertiary qualifications are more optimistic (57%) about AI than people with a high school education or less (29%) (Zhang & Dafoe 2019). Furthermore, tertiary students undertaking technical specialisations are more optimistic than students completing degrees in the humanities (Gherheș 2018).

Similarly, support for high-level machine intelligence is greater from people with computer science or programming experience (58%) than those without such experience (31%) (Zhang & Dafoe 2019). People's income also affects their attitude, with those earning over $100,000 annually, more optimistic about developing the technology (59%) than people earning less than $30,000 (33%) (Zhang & Dafoe 2019).

Finally, religion also influence's perception, with people identifying as not having a religious affiliation being more optimistic than Christian respondents (Zhang & Dafoe 2019). Furthermore, there is more optimism by people from developed countries than developing countries (Zhang & Dafoe 2019). Specific cultural contexts also account for such variation, such as Japanese support for the technology and the influence of Shintoism in the country, whereby there is no distinction between beings with and without a soul (as opposed to Western Christian notions of the soul and the afterlife). Japanese robots "share the same dignity as human beings and are widely accepted and integrated into daily life" (Righetti & Carradore 2019: 434).

Zhang and Dafoe (2019: 5) sum up these influences as "subgroups that are more vulnerable to workplace automation express less enthusiasm for developing AI", and due consideration should be given to these different demographic attitudes when studying perspectives.

3.5.2 Perceptions stemming from representations of AI

Existing research into audience perceptions of AI suggests that audiences are (at least partially) influenced by media representations of the technology, and this can be used as one lens to partially explain some of the general attitudes towards AI.

According to Nomura, Suzuki, Kanda, and Kato (2006), attitudes towards robots are more directly influenced by assumptions than by peoples' cultures and personal experiences, and cultural situations such as media and media distribution affect these assumptions. This might well be linked to emotional responses stemming from viewership as, according to Holt (2008), the subjects of science fiction focus abstract, emotionally relevant concerns, serving as objects

for emotional responses towards the subject matter. This is significant to consider since, according to McClelland (2016: 89) "mass media presentations of robots can shape wider societal attitudes towards real robots as they take their place in society".

Furthermore, the findings of Bristows' (2018: 21) study indicates that the public perception of AI seems to be more aligned to "the media hype cycle" than the view of experts. This hypothesis that public opinion of AI is at least partially influenced by the messages they receive from reporting on this area (Bristows 2018). Obozintsev (2018) also found that framing the technology in a positive manner whilst simultaneously addressing popular concerns is an effective strategy in promoting positive attitudes towards the technology.

This research suggests that media representations of AI do indeed influence audience perceptions. However, there has not been much research to this end. Furthermore, there has not been a study to date that considers the effect of fictional representations of AI in a particular text (or set of texts), through a comprehensive analysis of the meanings embedded within the text by considering all aspects of the medium, and the consequential audience perceptions of AI using audience research rather than examining (literal) textual cues.

This is problematic and points to a significant gap in the existing body of knowledge, as the cinematic medium allows for rich depth in meaning-making through, for instance, visual subtext, while "researching actual audiences rather than mere textual analysis is required to understand the complexities of the reception process" (Rauch 2018).

3.6 Conclusion

Fictional stories of AI (or thinking machines) have been told for over 2000 years. These fictional portrayals emerge in the form of science fiction, a genre that takes contemporary attitudes and debates and projects them into works of imagined futures. These portrayals also emerge through representing the technology most commonly through humanoid embodiment, and it is widely believed that while this is likely due to the need for character identification, it could also point to a future in which this form of AI gains more prominence.

If we categorise fictional representations of AI (or thinking machines) through the lens of identity exploration in terms of a parallel progression over time through technological innovation, these narratives reflect attitudes towards the technology in various eras of technological advancements.

Firstly, these representations reveal attitudes about divine intervention in identity formation and artificial creation before revolutionary industry. This shifted towards an exploration of what it means to be human during the first industrial revolution. This exploration was extended to actual real-world applications and considerations during the second industrial revolution. These explorations increased their frequency, level of complexity (due to the definition of AI), and themes concerning the possibility for the replacement of human identity during the third industrial revolution. Finally, the ongoing exponential technological advancement of the fourth industrial revolution has resulted in narratives exploring our newfound personal attachment and reliance on technology in all spheres of our daily life, in effect tracking the increasing integration of the technology (now with built-in and external sophisticated AI tools and applications) as it becomes *part of* our identity. Furthermore, throughout history, a recurring theme in fictional narratives depicting AI has been the loss of control.

Despite such widespread portrayals of the technology, research suggests that most people do not understand the technology, in that they are not aware that many of the applications and services they use regularly are embedded with AI. The public seems to believe that AI is still a futuristic technology due to their preconceived notions of what the technology might be, believing that common modern services and applications do not utilise this technology, as the AI in the technology is not at the level of advancements they believe 'AI' to be.

Research also suggests that people believe that contemporary AI has a low potential for significant benefits and risks presently and in the near future. However, there is a belief that this will increase in the years to come, especially regarding risks attached to the development of high-level artificial intelligence.

The variation of these perceptions is influenced by factors such as age, gender, level of education, industry, income, religion, and geographic location. The media also influences perceptions, contributing to identity formation and thus directly affecting perceptions whether through reinforcing pre-conceived ideas, or through the presentation of new ones. Existing research into audience perceptions of AI confirms the fact that audiences are (at least partially) influenced by media representations of the technology, and this can be used as one lens to partially explain some of the general attitudes towards AI. However, there has not been much research to confirm how this influences manifest as attitudes or expectations in practice.

Furthermore, there has not been a study to date that considers the effect of cinematic representations of AI in a particular text (or set of texts), through a comprehensive analysis of the meanings embedded within the text by considering all aspects of the film, and the consequential audience perceptions of AI using audience research rather than examining (literal) textual cues. This research output therefore bridges this gap in the body of knowledge. It draws on existing research, while adding to this knowledge by investigating both representation and affect, direct input and output, of a narrow field of texts. This allows for a more direct examination of the relationship between cinematic representations of AI and audience perceptions.

The following chapter establishes a theoretical framework after reviewing key theories and approaches necessary to measure, through analysis, portrayal and perception. This bridges the gap between the existing literature and the methodology of this particular study, while placing the analysis within certain parameters.

CHAPTER 4
THEORETICAL FRAMEWORK FOR PORTRAYAL AND PERCEPTION

4.1 Introduction

The previous two chapters provided an overview of the literature from a broad range of fields relevant to the solving of the research problem. This was achieved by creating a conceptual framework through which to understand AI in Chapter Two, and analysing how the notions in the conceptual framework have been postulated as being portrayed and perceived in Chapter Three.

This chapter acts as a bridge between the aforementioned conceptual and representational findings and the methodology of this particular study, by providing a theoretical framework after reviewing key theories and approaches necessary to measure, through analysis, portrayal and perception. Through a combination of conceptual underpinning, representational tradition and insights, and theoretical framing, the methodology of the study is appropriately constructed, framed, and limited.

Therefore, this chapter provides the background and framework for understanding portrayal and perception. Both of these areas require the definition and explanation of particular theoretical points of view that have been adopted and applied in this study, in their use as a set of tools and assumptions in the later analysis.

Firstly, this chapter examines the origins and foundations of semiotics by analysing the theoretical positions of structural semiotics, in order to highlight difference in social semiotic approaches. This is then transferred to cinematic social semiotic analysis particularly. Relatedly, and since cinema encompasses such a wide array of elements working together to create meaning, this chapter describes film theory positions on individual elements of a cinematic production, classed according to the categories of narrative, performance, medium aesthetics, and external factors.

Finally, theories on active viewership are outlined to provide a set of assumptions about the audience, acting as a lens through which to consider audience research and analyse responses. This therefore provides an account of assumptions made about audiences' decoding of media messages.

4.2 Social semiotics

Semiotics, at its most basic definition, is the 'science of signs' (Walliman 2011). The field of semiotics attempts to gain a deep understanding of textual meanings by interpreting elements of texts through signs (Walliman 2011; Stokes 2003; Sebeok 2001). Meanings are relational. We understand the meaning of something only in relation to the meaning of something else, whether through observable real-world relationships, convention, or a complex relationship between the two. Furthermore, even if there are observable real-world relationships, things still do not 'have' inherent meaning. Rather, these meaning exist in the human mind (Babbie 2012). The symbolic function of a sign, through representing or symbolising a real world entity or notion, conveys meaning through a process known as signification (Hall 2013).

A sign is any physical form that has been imagined, and through a physical medium, stands for an object, event, or feeling (Sebeok 2001). The word 'robot', an image of a robot, and the portrayal of compassion (for or from a robot) are all signs. These are known as referents, and a class of related objects, events or feelings is known as a referential domain (Sebeok 2001). Signs allow for the recognition of patterns, they serve as predictive guides for taking actions, and they serve as exemplars of specific kinds of phenomena (Sebeok 2001).

In fact, human intellectual life is based on the production, use and exchange of signs and representations (Sebeok 2001). We engage in sign-based representational behaviour when we talk, write, read, listen to music, watch films, or perform any other communicative act, allowing us to know, to behave, to plan, to socialise, and to communicate (Sebeok 2001).

4.2.1 Structural semiotics: Origins and foundation

Charles Sanders Peirce developed a triadic model (Figure 4.1) to explain the constitution of signs. His model consisted of a 'representamen', the form (also known as the sign vehicle); an object to which the sign refers ─ its referent; and an interpretenant, the 'sense' one makes of the relationship between the sign vehicle and the object and the effect it has (Chandler 2007: 30).

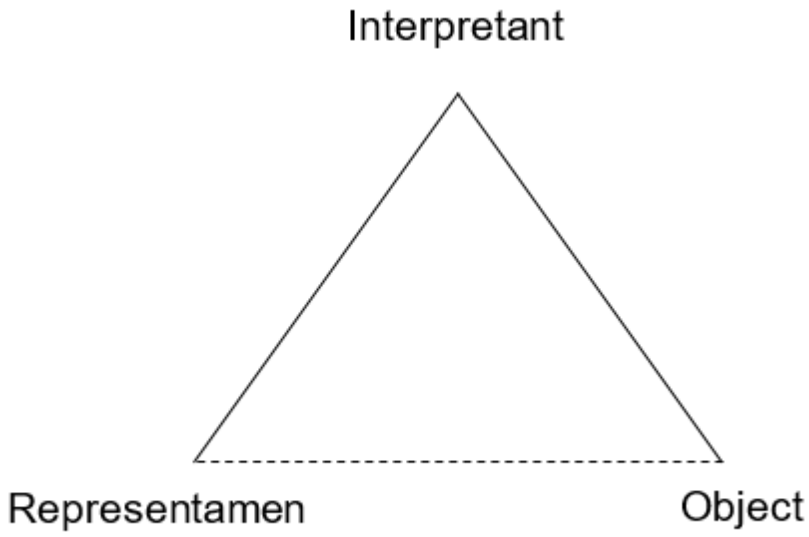


Figure 4.1: Peirce's semiotic triangle (Chandler 2007: 30)

Ferdinand de Saussure, on the other hand, explained signs in terms of a dyadic model (Figure 4.2) with the sign being the result of a signifier (the form) and a signified (the mental concept to which the sign refers) (Chandler 2007). The relationship between the two is referred to as 'signification', represented in Figure 4.2 by the two arrows (Chandler 2007: 14).

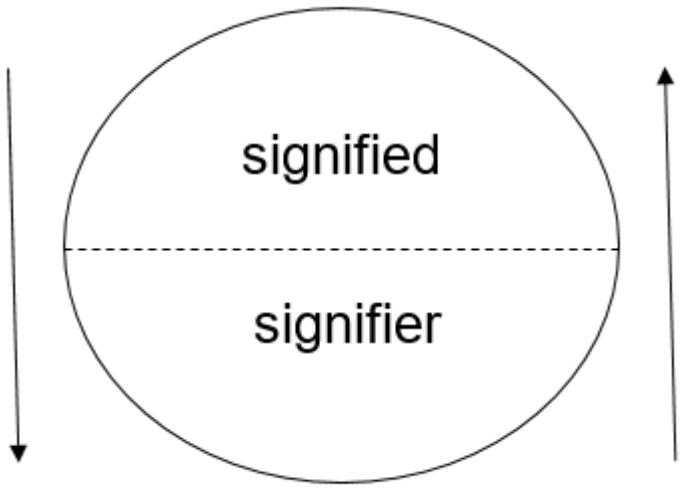


Figure 4.2: Saussure's model of the sign (Chandler 2007: 14)

Therefore, the Saussurean model describes the signified not as an external referent, but rather as a mental concept, and this is the key difference between these two models. The dotted line between the signifier and signified represents the fact that while these are essentially different

sides of the same coin (the sign), we can distinguish between these to explain the workings of signification. The signifier of a given sign can be classed under three modes of signification (Chandler 2007), depending on the relationship to its referent framed in terms of the constitution of the signifier/signified.

The symbolic mode points to a signifier/signified relationship that is arbitrary and based on convention (Chandler 2007; Sebeok 2001), such as the word 'robot'. The iconic mode points to a signifier perceived as resembling or imitating the signified (Chandler 2007; Sebeok 2001), such as an *image* of a robot. Finally, the indexical mode points to a signifier that has a physical or causal connection to the signified (Chandler 2007; Sebeok 2001), such as light beams on a robotic face, signifying eyes. This categorisation is not an absolute, however, and should be positioned on a sliding scale ranging from completely arbitrary, to complete resemblance.

The signified of a given sign consists of two parts, in the form of denotation and connotation. Denotation refers to the mental understanding of the literal meaning of the sign, while connotation refers to the socio-cultural and 'personal' associations of a sign (Chandler 2007). These associations might be related to a person's age, class gender, or a myriad of other factors ranging from emotional attitudes, to lived experiences, or ideological beliefs (Chandler 2007). The word 'robot', for instance, serves the purpose of signification by allowing interpreters to understand the literal meaning being conveyed, while creating a charge of emotion and attitude in the interpreter's mind (which might be positive or negative, depending on the aforementioned factors).

Figure 4.3 depicts the core system of subcategorisation in the signification process by means of a conceptual tree. Applying the word 'robot' to this conceptual tree, we can describe it as a sign, with a symbolic English word acting as the signifier. It has a signified which denotes an abstract group of entities in the world, meaning that the mental representation of such entities will differ according to the interpreter's knowledge and understanding thereof. In a similar manner, it has varied connotations based on personal, cultural, emotional, and ideological factors.

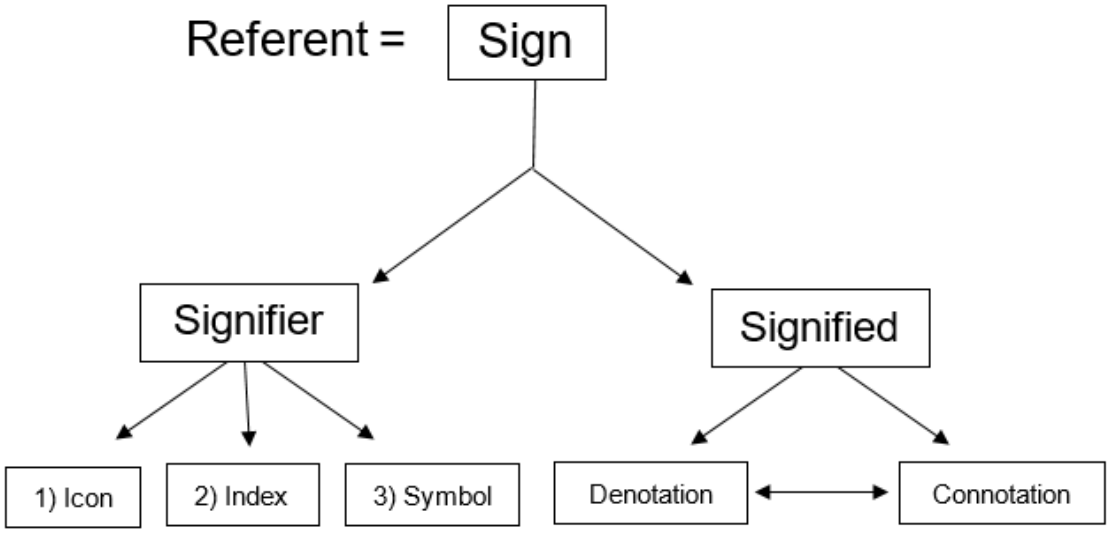


Figure 4.3: Conceptual tree of the signification process

Representational behaviour varies from culture to culture, so the signs that people use constitute a "mediating template in the worldview they come to have" (Sebeok 2001: 8). Words (or other types of signs) also shift their meanings (Hall 2013). Therefore, the signifieds attached to signs also change, shifting the conceptual map of the culture, allowing for different cultures at different moments in history to think about the world differently (Hall 2013).

Signs are interpreted through various codes, and each individual will learn codes specific to their socio-cultural positioning at various moments (Chandler 2007). Chandler (2007) groups these codes under three categories: social codes, textual codes, and interpretive codes. Social codes include verbal language, bodily codes, commodity codes, and behavioural codes. Textual codes include aesthetic codes, genre, rhetorical and stylistic codes, and mass media codes. Interpretive codes include perceptual codes, and ideological codes.

The creation of media texts requires the selection and combination of signs in relation to codes with which we are familiar, limiting the range of possible meanings they are likely to generate when read by an audience (Chandler 2007). Accordingly, in order to interpret a text, one needs social knowledge (knowledge of the world), textual knowledge (knowledge of the medium and genre), and knowledge about the relationship between the two (Chandler 2007), mediated through interpretive codes.

The act of explicitly highlighting the codes through which signs operate allows for a denaturalisation of signs, the deconstructing of a 'reality' (Chandler 2007). Barthes (1973) defined such naturalisation of ideology ─ constructing a reality in this regard ─ as myths. These myths serve the function (often by means of attaching certain connotations to certain signs in certain codes) of making dominant cultural values, attitudes and beliefs seen natural, timeless, obvious, common sense, thereby hiding the ideological function of signs (Chandler 2007).

Therefore, semiotics is an effective tool to uncover the nuanced meanings encoded in texts. However, pure Saussurian structuralist semiotics will not suffice if we are to consider the socio-cultural context of production. Structuralist semiotics focusses on formal systems rather than on processes of use and production (Chandler 2000). The emergent field of social semiotics, on the other hand, does account for the context of production. This is necessary since many films "conceal their context which is why it is important that analysis always situates both producer subjectivities and intentions and reception within their specific contexts" (Tomasseli 1996: 34).

4.2.2 Social semiotics: Key differences

According to Chandler (2000), there is nothing new about the social dimension of semiotics since semiotic codes are social conventions, and signs do not exist without interpreters. However, social semiotics as a field goes further than an increased emphasis on social codes, and these nuanced differences have led many scholars to consider the field as a separate branch of the field of semiotics.

Bezemer and Jewitt (2009: 421) define social semiotics as "the media of dissemination and the modes of communication that people use and develop to represent their understanding of the world and to shape power relations with others". Social semioticians see context as another set of texts, and it is similar to pragmatics in that it is concerned with meaning in context (Bezemer & Jewitt 2009).

Due to its social origin and the inevitable socio-political consequences of an emphasis on rules contained in logonomic systems (that is ─ rules prescribing the conditions for production and reception of meanings), social semiotics is invariably social in nature and in direct contrast with the "irreducible origin of structures in structural semiotics" (Vannini 2007: 117). Social

semiotics places emphasis on the participants of a semiotic activity as they interact in social contexts, instead of the structural semiotic emphasis on structures of signs and codes.

Furthermore, social semioticians emphasise logonomic systems as a social process, with people having different degrees of availability to power and instrumental resources (Vannini 2007). Social semioticians therefore dedicate attention to the effectiveness or practice of the power of logonomic systems on individual and group belief (Vannini 2007).

Social semioticians tend to be sceptical of dyadic models of the sign as inspired by Saussure, instead considering modified dyadic models or triadic models, often drawing inspiration from Peirce, who believed that meaning existed insofar as anything symbolic had practical consequences ─ with the interpretant gaining due consideration (Vannini 2007). Through this, social semioticians hold that signs do not stand for something pre-given that transcends use, instead they work as 'semiotic resources', shaped by the way in which people use them to make meaning (Vannini 2007).

The notion of semiotic resources is therefore central to social semiotics. Semiotic resources are the actions and artefacts that are used to communicate (van Leeuwen 2005). These might be produced physiologically with, for instance, our vocal apparatus, or through technologies (for instance, with pen and paper, or computer hardware and software) (van Leeuwen 2005). The term is preferred over 'signs' in social semiotic circles, due to the belief that signifying systems are not a set of rules but instead resources for making meanings (van Leeuwen 2005).

Resources are said to have a theoretical semiotic potential and an actual semiotic potential, each describing the potential for the resource to agentically achieve a communicative goal. The former consists of all its past uses and potential future uses, and the latter consists of the uses that are known by specific users with specific needs in specific contexts (van Leeuwen 2005). An analysis of the semiotic potential of a resource is an analysis into the manner in which semiotic resources are used in specific historical, cultural and institutional contexts (van Leeuwen 2005).

However, this is not to say that the notion of the 'sign' is excluded from social semiotics, but rather that the process of signification is subject to the "*interest* of sign-makers, their *availability* of semiotic resources and the *aptness* of those resources to the meanings which they wish to realize" (Bezemer & Jewitt 2009: 423 - 424) [emphasis in the original]. Bezemer

and Jewitt (2009) describe signification in this regard as people bringing together semiotic resources (signifiers) with the meaning (signifieds) that they wish to express.

This is mediated through discourses, which are socially constructed knowledge of some aspect of reality, used as frameworks to make sense of things (van Leeuwen 2005). Discourses are "system[s] of meaning which [are] confined to a particular group of people in a specific area of interest doing specific ideological work" (Tomaselli 1996: 40). Through this process, knowledge is constructed.

Therefore, the operation of discourse in social semiotics is akin to the operation of codes in structural semiotics. Codes can be seen as developing agreed upon meanings, thus promoting or opposing a dominant ideology or worldview (Tomasseli 1996). The key difference between the two is that the discursive conditions under which resources must operate are seen not as fixed or natural, but social rules in which the texts operate. Therefore, descriptors of codes from the vast tradition of structural semiotics may provide insights into discourses as a point of departure, but these should not be seen as fixed, nor that meaning cannot escape these limitations. Rather, these are social rules to help facilitate meaning.

According to van Leeuwen (2005), evidence for the existence of a given discourse comes from texts, specifically from the similarity between the things that are said and written in different texts about the same aspect of reality. Van Leeuwen (2005) states that fictional representations bring difficulty in linking discourses and real social practices and actors, as they transform these real-world entities by locating them in a mythical or distant past or future to allow them represent more than one set of social actors and social practices at the same time. They also represent key themes relevant to a wide range of contemporary social practices (van Leeuwen 2005). Genre acts as a vehicle to study the manner in which semiotic resources are used to enact communicative interactions, with the notion of style integral to studying how semiotic resources 'perform' genres (van Leeuwen 2005).

Kress and van Leeuwen (2006) describe the communicative system as serving three meta-functions simultaneously. The ideational function constructs representations of the world, the interpersonal function enacts or helps to enact interactions (characterised by specific social relations and social purposes) and the textual function projects communicative acts into larger wholes with communicative events or texts realising specific social practices (Kress & van Leeuwen 2006). Kress and van Leeuwen (2006) have used this model to describe various

resources of visual communication, including composition, gaze, angle, size of frame, and colour, in making the argument that these resources are part of the grammatical system of images.

The concept of modality is also central in social semiotics as a tool to study the use of semiotic resources in creating the truth or reality values of their representations (van Leeuwen 2005). Modes are socially or culturally shaped resources for making meaning (Bezemer & Jewitt 2009). In order to determine whether a set of resources (such as colour) constitutes a mode, it should be able to satisfy all three of the aforementioned meta-functional criteria. The relationship between modes within texts might realise tensions between the aspects of meaning in a text ─ being meaningful in and of itself (Bezemer & Jewitt 2009).

These modes enter into coding orientations informing the way in which texts are coded within specific contexts or by specific social groups (Kress & van Leeuwen 2006). These coding orientations include four broad categories. Technological coding orientations results in the audience evaluating modality based on the representation's effectiveness as a 'blueprint' for the user (Kress & van Leeuwen 2006). Sensory coding orientations, used in contexts in which the pleasure principle reigns, sees modality evaluated upon the ability of the image to awaken sensory responses (Kress & van Leeuwen 2006). Abstract coding orientations sees modality evaluated based on the basis of its fidelity towards abstract ideas or essential qualities of phenomena (Kress & van Leeuwen 2006). Finally, naturalistic coding orientations, typically found in photographs (and video), measures the correspondence between the object of representation and how it would be viewed by the naked eye (Kress & van Leeuwen 2006).

Coding orientations mean that modality markers in isolation are not sufficient to evaluate truth value or credibility, but that purpose and context are necessary factors. For example, a high degree of colour saturation may mark low modality in the naturalistic orientation, but a higher one in sensory or abstract texts (Kress & van Leeuwen 2006). Furthermore, an audience member might utilise different coding orientations at different moments, such as through non-naturalistic orientations in activities of group allegiance, reverting to naturalistic orientations at an individual level (Kress & van Leeuwen 2006).

Accordingly, social semiotics encourages an identification of the modes, the resources of those modes, the context of those modes, and the meaning makers, to understand how sign makers exploit the potentials of resources to create particular meanings (Bezemer & Jewitt 2009).

4.2.3 Cinematic social semiotics

Perhaps due to its relatively recent origins, social semiotics provides more tools for cinematic analysis than structural semiotics (with many seminal structural semiotic works naturally emphasising written texts). However, it is still important to look beyond 'pure' social semiotic writing as many other scholars (indeed many making reference to the principles of social semiotics) have established additional fundamental considerations to facilitate the task of cinematic semiotic analysis.

Contemporary semioticians do not study signs in isolation, but rather through a sign-system such as a particular medium or genre (Chandler 2007). By the act of analysing a representation within a particular medium, we might find inherent significant considerations to the overarching nuances of signification within that medium. For instance, audio-visual media are regarded as 'more real' than other forms of representation (Chandler 2007), thereby already finding themselves more skewed towards a naturalistic coding orientation and susceptible to greater ideological myth.[1] A transference of signification from, for instance, a book to a film, might greatly influence a signified (Chandler 2007), by the mere fact of it being a film and operating within such textual codes.

A film's syntagma is the way in which individual frames of picture and sound have been combined to produce certain meanings (Tomasseli 1996). Unlike written and spoken communication, the filmic medium not only considers linear constructions of language, but also the composition of space and time, through the *mise-en-scène* and *montage* (Monaco 2009).

According to Sebeok (2001), films usually involve at least four codes, one visual, and three auditory (speech, music, and sound effects). These are of course medium textual codes, but films embed additional generic textual codes, social codes from which to construct meaning in performance, and interpretive codes from which to perceive these meanings mediated through particular ideologies. The nature of the filmic medium means that viewership depends on the operation of multiple codes from multiple categories, as well as a deciphering of multiple forms of modality.

[1] Although, this is in terms of an overarching view of coding orientations for the text. Isolated elements might be subject to different orientations depending on the context and purpose within the genre and medium.

The ‘meaning’ of a specific film, torn “internally (structurally) between the system as culture and the syntagm as nature” (Barthes 1977: 51), is produced under the conditions of a process known as abstraction. Abstraction involves a simplification of real-life experiences within narratives, with elements only fundamental to the goals of the characters forming part of the plot, entering into the meaning-making framework under these conditions (Mar & Oatley 2008). As abstractions, films provide explanations of what goes on beneath the surface to generate observable behaviour, suggesting a direct relationship with ‘real’ behaviour (Mar & Oatley 2008).

Referential meaning is created when a film alludes to knowledge outside of the film which the viewer is expected to recognise (Bordwell & Thompson 2008). Naturally, there will also be a range of explicit and implicit meanings within a film, with the former concerning the presentation of overt significance, and the latter concealing this significance, for discovery upon analysis or reflection (Bordwell & Thompson 2008). Finally, films also are subject to symptomatic meaning, wherein significance is divulged by virtue of its historical or social context (Bordwell & Thompson 2008).

Monaco (2009) states that while iconic and symbolic signs are clearly present in cinema, indexical signs require a closer examination, as it may not be possible to understand what the sign systems denote without an attached connotation. Much of the connotative power of a film depends on indexical devices (Monaco 2009). The connotative power of films (while very often indexical, but not limited to this category of signification) often materialises in ‘tropes’. Tropes are the rearrangement of a sign’s signifier and signified to create a new relationship between the two (Monaco 2009).

Metaphor involves the comparison of one thing to another, thereby operating in either the iconic of symbolic mode of signification (Chandler 2007). Many metaphors become normalised after habitual employment, to the extent that viewers do not even realise they are engaging in a metaphoric comparison (Chandler 2007). In film, a pair of consecutive shots is metaphorical if there is an implied comparison between the shots (Chandler 2007).

Visual metaphors serve a function of ‘transference’, transferring certain qualities from one sign to another, and might relate to spatial organisation, ontological associations of entities and substances with emotions, ideas and activities, or structural metaphors involving conceptual linkage (Chandler 2007).

Irony operates similarly to metaphor, in that the signifier of the ironic sign seems to signify an aspect of reality, but we know from another signifier that it actually signifies something very different (Chandler 2007). While metaphor highlights similarity, irony highlights difference. Irony may reflect the opposite of the 'truth' about external reality, or thoughts or feelings of the speaker or writer (Chandler 2007). This may manifest as dramatic irony, wherein the audience knows more about events than the characters do, situational irony, where actions have an opposite effect as to what was intended, and verbal irony, in which a speaker's intentions are opposite to what they are saying.

On the other hand, metonyms involve a comparison in which an associated detail or notion is used to invoke an idea or represent an object (Monaco 2009), with the signifieds having a direct or close association to each other. The filmic medium allows for rich statements in this regard, as a result of the compression of associated details within the limits of a frame, and thus metonymy has become a sort of cinematic shorthand (Monaco 2009). These associations are indexical in nature (Monaco 2009; Chandler 2007), and are marked by a depicted object representing a related but non-depicted object or idea (Chandler 2007).

Finally, synecdoche represents the substitution of a part for a whole, a genus for a species or vice versa (Chandler 2007), thereby connecting an object to another object or feeling and invoking related connotations in the mind of the receiver. It may involve a part standing for a whole ('wheels' to describe a car), or a whole for a part (the law standing for police) (Chandler 2007). It can further be divided into hypernymy and hyponymy, with the former describing the use of a hyponym for a superordinate class which includes it (such as a 'Hoover' for vacuum-cleaner), and the latter describing a superordinate for a hyponym (such as a machine for computer) (Chandler 2007).

The frame of a visual image itself acts as a synecdoche in that it suggests that what is portrayed is a 'slice-of-life', and that the narrative world outside the frame is carrying on in the same manner as the world depicted within it (Chandler 2007). In fact, there is the operation of a synecdoche during any attempt to represent reality, reflecting the most direct link of signification in audio-visual media, since it is an integral part of its being (Chandler 2007). Accordingly, the sense we make of cinematic connotations depends on comparisons of an image syntagmatically (with images that came before and after), and paradigmatically (with images that were not chosen), and therefore our sense of cultural connotations depends on understood comparisons through synecdoche and metonym (Monaco 2009).

Connotations can be further engineered in representations by changing the form of the signifier while keeping the literal signified, such as the choice of words (robot or machine), or subtly changing the style or tone such as choosing a sharp or soft focus (Chandler 2007). These additional connotative considerations also need to be analysed when conducting a cinematic social semiotic analysis, along with denotative meaning, all the while considering the various codes in operation.

4.3 Film theory and meaning

It is clear then, that the 'meanings' produced in the selected texts are not singular, nor can the identification of meaning be a linear exercise. Instead, each text has multiple meanings, expressed through multiple forms, signifying multiple aspects of reality, using multiple (and varied) codes. We might be able to deduce overall messages from this multiplicity of meanings, but this requires an analysis all of the various cinematic elements, or modes. For a feature film, this is naturally a lengthy process.

It would perhaps be beneficial to isolate a specific element of the film (for instance, the performance of a character portraying a robot), and analyse this to isolate meanings particular to this aspect of the film. However, in reality, audiences do not only consider 'pure' performance through viewership. All aspects of the film, with all of its intricate modes of delivery (due to the nature of audio-visual media), are decoded by an audience either consciously or subconsciously (of course dependant on a multiplicity of factors), to produce meanings in certain contexts. Therefore, the structure of such an analysis is a complicated endeavour, and it is necessary to define a clear structure for the purposes of analysis.

Van Leeuwen (2005) outlines an interlinked system of four dimensions that form part of every semiotic artefact. This model includes rhythm for coherence and structure over time, composition for coherence and meaningful structure to special arrangements, information linking to provide cognitive links between the items of information in time and space, and dialogue to analyse the use of audio in highlighting the relationships between modes (van Leeuwen 2005).

Chandler (2000) categorises several key considerations when conducting a semiotic analysis, including: Textual identification; purpose of analysis; important signifiers and what they signify; modality of the text (truth); paradigmatic analysis (medium, generic, and thematic considerations); syntagmatic structure; rhetorical tropes; intertextuality; semiotic codes; social

considerations; and benefits of the semiotic analysis. As part of the paradigmatic and syntagmatic interrogations forming part of a more general semiotic analysis, Chandler (2007) outlines six textual features typically listed by film theorists including narrative, characterisation, basic themes, setting, iconography, and filmic techniques. These are all also under the influence of particular generic codes.

Codes also need to be analysed in relation to other codes, as well as to intertextual influences on such codes (Chandler 2007). Furthermore, Monaco (2009), in considering traditional examinations in visual art of colour, line and form, explains that a framed image has two important considerations, namely the limitations that the frame imposes, and the composition of the image in the frame.

As will become apparent in the following chapter, it is necessary for this particular study to consider semiotic markers in relation to aspects of cinema as a medium, rather than using cinematic medium markers or broad categories (such as rhythm) to highlight pre-established semiotic modes of meaning. A use of visual metonymy, for instance, should not be analysed by considering aspects such as editing, visual effects and sound. Rather, editing, visual effects and sound should be marked as categories of meaning, forming part of a complex whole, which allow for the transmission of a visual metonym. This fosters a mode of analysis that is comparable as part of a larger analysis for a particular purpose, and in which particular aspects of the medium are not susceptible to neglect, such as considering sound and editing but neglecting visual effects (which might have elements pointing to an opposite metonymic meaning, thus being meaningful in and of itself).

As a baseline linkage system to structure analysis, encompassing the modes and semiotic resources of those modes in the spirit of social semiotics (Bezemer & Jewitt 2009) the categories of Passchier's (2007) Entertainment Value Assessment Matrix (EVAM) are useful to consider. This model was designed to measure the entertainment value of film, television, and live performance productions in the pre-production phase, and the categories within this model act as a unified whole of the production of a creative output.

The EVAM categorises an entertainment production into five components, namely narrative, performance, medium, aesthetics, and control. Narrative refers to elements working towards audience engagement with the production, while performance refers to elements working towards emotion within a production (Passchier 2007). Medium refers to the delivery of the

narrative and performance elements through the cinematic medium (Passchier 2007). Similarly, aesthetics refers to the design and fabrication of visual form to amplify the narrative meaning (Passchier 2007). Finally, control refers to elements in need of consideration for economic viability and sustainability in the product market (Passchier 2007).

The model was specifically designed to measure the entertainment value of productions, and therefore using the model for analysis does require some restructuring. This includes the addition of notable dialogical moments in narrative, notable behavioural codes in performance, and the restructuring of 'control'. Since this category refers to both internal and external factors, this is repurposed to only include general and specific relevance as an external consideration, with marketing and producing considerations omitted.

However, the model still provides an appropriate linkage system to compartmentalise particular aspects of cinema. The categories of narrative, performance, medium, aesthetics and external factors can therefore be utilised to group elements (modes and their resources) frequently used in cinematic semiotic analyses, using film theory for a theoretical overview of the deployment of meaning through these elements.

4.3.1 Narrative

Narrative elements include conceptual themes, plot, conceptual relevance, notable dialogue, conflict, point of view, and setting. Bordwell and Thompson (2018) define themes as the broad concepts emanating from the abstract quality of implicit meanings. These are the 'lessons' a film divulges, the exploration of general unifying central concepts acting as a foundation for a text.

Themes encompass more than just an abstract concept such as 'war', or even 'the effects of war', and rather explain specific meanings to be uncovered in the entire filmic system, such as 'in times of war, one might be surprised at what one might do to survive'. The concrete realisation of a film usually stems from conceptually thematic exploration, yet as Bordwell and Thompson (2008) note, a film's thematic meanings should be interpreted only in relation to the film's total system. Therefore, while all cinematic aspects emanate from themes in practice, themes can only be understood by considering all cinematic aspects in analysis.

An initial reading of a text might give clues or hypotheses as to the themes in a film, but this should be tested in relation to an analysis of all other elements. From this, we can alter the

initial thematic explanation based on later findings, while still having a broad lens through which to view potential significations of meaning. Conceptual themes, then, can only be fully realised if read in relation to other cinematic aspects.

The plot of a narrative refers to the specific, pre-determined organisation of narrative events in order to ensure narrative comprehension, achieve abstraction, and create engagement, conflict, and identification. An analysis of a plot will inevitably reveal the 'story', but it also serves a more important function in creating an environment for the other nuanced elements of narrative meaning to emerge. In other words, an outline of the plot reveals much of the denotative meaning of a film, while paving the way for the interpretation of resultant connotative meaning through both structural and non-structural elements.

Both of the selected texts conform to a linear three-act plot structure. This is unsurprising because, as Smith (2000: 88) notes, this plot structure has become the "trusty template that defines American cinema". By conforming to this structure, the selected films are able to build towards a gripping climax (Aronson 2010). They also encourage the audience to empathise with a single protagonist, as their linearity allows the audience to "put [themselves] in the situation of the other, and ... bring home to [themselves] every little circumstance of distress which can possibly occur to the sufferer" (Smith 1759, cited in Mar & Oatley 2008: 180). The three-act structure allows for abstraction to take the form of events. The length of these events, based on their importance to the *fabula*, simultaneously creates the notion of importance for such events due to the amount of screen time (Bal 2017).

The three-act structure has five basic elements. The first act is the establishment of narrative world. A plot/turning point then causes conflict, triggering the second act, with a confrontation of that conflict. A second plot/turning point further complicates matters in this confrontation, triggering the third act in which the climax of the film takes place, and the narrative is consequentially resolved in the third act.

Many of the additional elements used to define narrative events taking place within this structure draw from the work of Vladimir Propp and Tzvetan Todorov. Propp (1968) outlines 32 narrative functions, divided into six sections: Preparation, complication, transference, struggle, return, recognition. On the other hand, Todorov[2] (1977) outlines narratives as

[2] See section 3.2 for a discussion on science fiction genre conventions in relation to Todorovian narrative categories.

beginning with an equilibrium where potentially opposing forces are in balance. This is then disrupted by some event, usually through the actions of a villain, setting in motion a quest towards (and reaching) a second yet different (and preferably more stable) outcome.

The EVAM model expands the three-act structure over ten distinct beats in a process referred to as Explosive Theme Bundling (ETB). Rather than acts, the model describes events in terms of Problem-Solving Priority Orders (PSPO's), which are based on the internal mechanisms people use to navigate life through managing solutions to problems (Passchier 2007). This is due to the model's emphasis on engagement and emotion through problem solving activities, and the impacts of problem-solution relations in cinema (through conflict) on a viewer's own internalised strategies thereof.

Accordingly, acts one, two and three become the Established Problem-Solving Priority Order (EPSPO), the New Problem-Solving Priority Order (NPSPO), and the Final Problem Solving Priority Order (FPSPO) respectively. The EPSPO (beat one) provides a 'set up' of the narrative world, characters and the usual way of life. This PSPO is disrupted by the First Explosion (beat two, plot point one), an event making the previous way of life inappropriate or inaccessible, thus causing conflict. Resultantly, there is a Problem Solving Vacuum (PSV) providing a view into the emotion attached to this point of crisis (beat three). From this point, the character rises to the occasion during the Problem Solving Attraction (PSA) phase, allowing for identification and empathy (beat four). Potential solutions available to the character arise (beat five) through Problem Solving Digits (PSDs), and the character selects one of these as their solution, acting on it (beat six). There is a resultant confrontation of the problem in the NPSPO (beat seven), which is ultimately proven to provide a false sense of security as unforeseen problems arise in a Second Explosion and its vacuum (beat eight, plot point two). Now with fewer resources than before, the character must (in an all or nothing moment) confront the conflict and antagonising force in the Climax (beat nine), resulting in the resolution or FPSPO (beat ten) where life is now even more stable than during the EPSPO.

Following on from the ETB structure, the EVAM also introduces the notion of conceptual relevance. This refers to how an organism responds to an event problem in the narrative (Passchier 2007). This problem arises from the environment, and yet it is with the limited resources within the environment that the character has to solve the problem (Passchier 2007). This solution is internalised as a PSPO, and if this problem-solution activity is relevant to the audience, it ensures audience engagement. Therefore, through outlining the particular

relevance of the solving of the central problem in the narrative, we are able to uncover the means for engagement as part of a strategy to measure meaning creation holistically. In other words, this allows for an examination of meaning in relation to problem-solution internalisation.

While film is a visual medium, and technically a film can provide rich meaning without any dialogue whatsoever, an analysis of key dialogical moments and the meanings imparted from this is valuable in supporting the rest of the narrative meaning in the film.[3] When considering dialogue as part of analysis, notable meanings about (for instance) character attitudes and express conflict, either on the surface or subtextually, can be uncovered. To limit the subjectivity of making a value judgement of what might be considered 'notable', this element should be limited to an examination of dialogue that may add or subtract from meaning in the previous narrative elements (particularly plot).

This may also uncover moments that either show isolated incidents of (either supporting or opposing) character attitudes towards the subject, ethical concerns about the technology more generally, or optimism about the benefits or function about the technology. However, these isolated dialogical instances should be viewed in relation to the unfolding of events to highlight additional meanings (such as irony).

The conflict within a narrative might well point to societal conflict, as many studies have attempted to examine.[4] As such, the conflict within a narrative requires careful consideration. Conflict stems from the disequilibrium caused in the first plot point (explosion) by the antagonising force (not necessarily a character), and is based on binary opposition (such as good and evil, security and uncertainty). Since conceptual relevance concerns the solving of an event problem, that problem is caused by the conflict, and thus the solving of the conflict is also meaningful to an audience.

Conflict may be internal, involving a battle between a character and themselves, such as making a decision, discovering their identity, and so on. It may also be external, such as a battle between a character and another character, or nature, or society more generally. This 'battle' takes the form of the rising action throughout the plot, ultimately reaching a crescendo during

[3] These meaning are likely to be denotative but might potentially also be connotative through, for instance, subtext.
[4] For instance, as previously highlighted, Arnold (1998) argues that the conflict in the first two films of *The Terminator* series refers to the conflict with automation in the socio-cultural context of production.

the climax. According to McKee (1997), one ought to describe conflict as occupying three levels. This is achieved by examining firstly, who or what is driving the story, as well as their motivation; secondly, who or what is blocking them or it; and thirdly, what this antagonistic force desires.

Point of view refers to the perspective from which an audience views the unfolding of events. This allows for a view into the dispersion of meaning through dialogue and character actions, as the particular perspective employed might allow for greater understanding of the meanings produced. For instance, selecting the perspective of a particular character as an event unfolds might suggest that that character is the one with which the audience should identify, altering the meaning of that event in that regard.

According to Bordwell and Thompson (2008), the point of view or mode of address employed in a text depends on the textual context (convention of the genre and syntagmatic structure), social context (external to the text such as economic factors and composition of the audience) and technological constraints (such as features of the particular medium).

Films are usually told from a detached, third-person omniscient point of view (Bordwell & Thompson 2008; Monaco 2009). However, this overarching mode of address may shift at various moments over the course of the film. For instance, camera work or voice-overs may provide a particular view of a character's experience of the diegetic world at a particular moment, and this may even shift between multiple characters producing polyvocality (Bordwell & Thompson 2008).

The final narrative element, setting, refers to the world of story. Setting refers not only to what is seen in the frame at a given moment, but through the suspension of disbelief, what is perceived to exist within that narrative universe ─ the laws governing its existence. Accordingly, a narrative needs to obey its own internal laws of probability, limited in the possibilities and probabilities of the narrative world (McKee 1997).

Naturally, an examination of the narrative world itself as an entity might well uncover a range of unique meanings. This is particularly useful for this study, since in examining the meanings of representation of AI in an imagined future, it is imperative that due importance is placed on the imagined future itself and not just the events taking place in that future. McKee (1997) outlines four elements that comprise a setting, including the period, or the narrative's place in time (past, future, present); the duration, or the narrative's length through time; the location, or

narrative's place in space; and the level of conflict, or the narrative's position on the hierarchy of human struggles (either internal, interpersonal, institutional, or environmental).

4.3.2 Performance

Performance elements include character identity, emotional relevance, social codes, and character arc. Character identity concerns the construction of the fictional characters within the narrative. As such, this involves an interrogation of the nuanced aspects that make up the characters in the selected texts in terms of character identity.

Character identity is to performance what plot is to narrative. In other words, this is an important point of establishment for other performance considerations. The selection of key characters naturally arises from narrative considerations, whereby a protagonist and antagonising force (be it a character or some other force) can be identified. From this, central characters can be highlighted in relation to their function within the narrative. Characters are often described in terms of certain roles in fulfilling the narrative in this manner. For instance, Propp (1968) defines several characters to this end, including the hero, the villain, the donor, the helper, the princess, the dispatcher, and the false hero. There are various other descriptions of these character-narrative roles (archetypes), but an in-depth analysis of these roles is not necessary for this study.

Fiske (1987) states that a single character can occupy several roles or 'spheres of action'. A dramatic character is a personality with a character, a simplification of (States 1985), or a metaphor for (McKee 1997), human nature. Identity, on the other hand, attaches the person, character, and personality to the world, thus creating a situation in which conflict might arise (States 1985). A focus on character identity particularly, rather than personality or mere characterisation, enables an examination of character beyond a brief outline of traits and instead providing a window into the complex, multiple identities at different moments, which is a key aspect of character design in cinema (Leal 2017).

According to Bal (2017), the major attraction of a narrative is the so-called 'character-effect', which occurs when an audience identifies with or against a character due to a resemblance between humans and fabricated figures. This can only be achieved once the audience receives a clear indication of the character identity.

To this end, McKee (1997) states that there are two elements comprising the psychological design of a character. Firstly, characterisation refers to the observable qualities that make the character unique, and secondly, true character reveals the person 'behind the mask', and this is revealed when a character expresses a choice through a dilemma (McKee 1997). The notion of 'dimension' in character occurs when there is a contradiction of true character, or between characterisation and true character (McKee 1997).

The translation of this 'character-effect' might seem incompatible with non-human (robotic) characters at first glance. However, according to Leal (2017), the android's gaze provides the energy to empower the entire character, to the extent that only a single eye as a red dot is needed to personify certain characters (such as HAL in *2001: A Space Odyssey* 1968), thus achieving relatability and a character-effect.

According to Leal (2017), the depiction of android characters provides an explicitly conscious account of cinema's violation of the assignment of a singular identity to each exterior appearance. In other words, due to the detachment of identity from human character, mediated through forms and techniques amplifying a particular gaze, the multiplicity of identities in android characters can be explored, and in doing so highlight how 'disturbing' mask wearing through performance might actually be (Leal 2017).

Emotional relevance refers to a character's emotional and behavioural adjustment in order to generate solutions and implement them effectively (Passchier 2007). This adjustment is relevant to an audience if the behavioural change is socially usable, whether it can be internalised by the audience as effective and aspirational, or as a better coping skill to solve the particular emotional problem (Passchier 2007). Therefore, through linking character identities to plot, conceptual relevance and conflict, we can uncover the emotionally relevant core through the engagement of the protagonist's problem-solving mechanisms, and the meaning this might have for an audience.

Character arc can be seen as an expansion of emotional relevance, in that it maps behavioural change of characters through the narrative. However, while emotional relevance operates in relation to what might be internalised by an audience, character arc is more generally related to the entire structure of that character changing emotionally within the progression of the story.

Finally, notable social codes function in a similar manner as notable dialogue in narrative, in that it provides additional meaning-making resources that might be overlooked by purely

focusing on character itself. This allows for meaning to be uncovered from character interactions through performance ─ whether interpersonally, in isolation, or environmentally ─ and the additional meaning making functions such interactions might add to the system of meaning.

Therefore, the realisation of dialogue and plot through performance might support or oppose the meanings thereof (such as through the use of tone, or other verbal language sub-codes). However, this also includes bodily codes such as contact, proximity, expression, gaze, contact, gestures, and posture, based on the function of the performative moment in relation to the overall system.

Smith's (1994) character empathy model offers an approach for the examination of narrative and aesthetic cues in constructing character identification and empathy. The model proposes three stages in the process, and is particularly useful for analysing both texts in relation to each other. The first stage involves a recognition of a character, including all quasi-extensions of a human body such as face, eyes, and wardrobe (Smith 1994). This ensures that a viewer receptive to alignment, which is the second phase of this model. This is based on medium and aesthetic manipulations influencing a viewer's access to a character spatially, as well as narrative manipulation allowing access to their psychological state (Smith 1994). The final phase, allegiance, concerns a moral and ideological evaluation of the character on screen. This model provides a suitable method to account for audience identification and empathy with a humanoid AI character.

4.3.3 Medium

Medium elements include cinematography, visual effects, sound design, and editing. The primary considerations for the purposes of analysis of cinematography in this study concerns framing, with a particular emphasis on shot size, camera angle, and camera movement. Each of these considerations serve to divulge particular meanings. This is due to the fact that in visual representation, social distance is related in part to apparent proximity (Bordwell & Thompson 2008).

However, it would be incredibly time-consuming to analyse each individual shot in each film. Rather, key trends of the delivery of certain characters or certain moments provide insights as to the efficacy in supporting or contradicting conceptual themes. This is also particularly useful as the particular meaning of a camera treatment of a shot in a scene might not be able to be

declared as an absolute (Bordwell & Thompson 2008), but this meaning becomes clearer if viewed in relation to the narrative and performance considerations, as per the EVAM model.

Since camerawork involves a depiction of proximity to characters, naturally certain shot sizes might divulge certain meanings. Shot sizes reflect degrees of formality, with close-ups signifying intimate or personal modes, and long shots signifying impersonal modes (Bordwell & Thompson 2008). Furthermore, according to Leal (2017: 155), "the close-up renewed the reputation of the face and eyes as bearers of an individual's true self". Similarly, the angle of view might, for instance, represent a character as small or insignificant using a high angle (looking down on the subject), or superior using a low angle (Bordwell & Thompson 2008).

Of course, rather than a straight reading of camerawork and the effects of proximity, a close-up in context may also, for instance, produce an uncomfortable intrusion into a character's space, rather than a warm, welcoming entrance thereto. Close-ups may even deprive the viewer of setting and create a claustrophobic disorientation (Monaco 2009). Juxtapositions of meaning may therefore also be uncovered in this regard.

Since science fiction relies so heavily on computer-generated effects, the use of visual effects in science fiction cinema also serves many of the aesthetic elements such as the physical manifestation of the gaze as outlined in performance. However, it is indeed a use of the cinematic medium, technically speaking (and technically produced), and the use of this technology itself is an important consideration.

Therefore, the description of visual effects should be viewed in relation to elements enhancing and delivering the narrative and performance, not in terms of form, but in terms of the link between dramatic function and external influence. This is then primarily to analyse, rather holistically, the general use of visual effects in the films, from Telotte's (2001) point of view of science fiction cinema highlighting a rather complex relationship to technology due to its reliance thereof.

Furthermore, sound can actively shape the perception and interpretation of an image (Bordwell and Thompson 2008). Indeed, the use of sound may anticipate another element, or relay attention to it (Bordwell & Thompson 2008). The description of sound and music as an enhancement of narrative and performance should be realised similarly to cinematography. Elements such as music, sound effects, foley, and ambience should be analysed in relation to previously established meanings, with a particular attention as to whether the sound is diegetic

or non-diegetic, and parallel (synchronous) or contrapuntal (commentative) (Monaco 2009), and how the use of the sound enhances and delivers such meanings more generally.

Finally, editing techniques have the ability to alter the meanings of a film dramatically.[5] Imagine, for instance, that a viewer was presented with a chronological list of shots in relation to the moment in the plot, which they were able to select and view at various moments. This would do away with the particular timing and rhythm of a scene, and it would afford the viewer a more active role in deciding on the importance of certain elements (due to their selection and screen-time) in a scene. A cut to a close-up, for instance, might direct a purposive shift in attention (Bordwell & Thompson 2008).

Therefore, this element serves a variety of functions in relation to meanings already established, as well as new meanings through its execution. Meanings can be uncovered by analysing the overall edit pattern at various moments in the plot, screen time afforded to particular characters, timing of shots at particular moments, the effects of certain types of cuts (as already outlined in terms of visual metaphor), and the manipulation of the progression of time in creating or enhancing meaning.

4.3.4 Aesthetics

Aesthetic elements include colour, character styling, location, props and iconography, and location and lighting. According to Adorno (1997), meaning is only legitimate in an artwork if it is objectively more than the work's own meaning, and this meaning is produced and reproduced aesthetically. Much of this aesthetic divulgence of meaning is through the use of colour.

According to Bellantoni (2015), colour is used in cinema to subliminally layer a story, and a slight variation in colour might have a profound influence on behaviour. Different films use colour in different ways, such as to support the definition or evolution of characters, and expand the story (Bellantoni 2015). Therefore, it is important to note the use of colour in conjunction with other meanings. This can be achieved through isolating the use of a single colour throughout a film or in particular moments, or analysing the use of colour harmonies to this

[5] Additionally, this medium aspect distinguishes cinema from other visual arts.

end. The use of these different harmonies might signify, for instance, similarity, opposition, conflict, and complication in context.

Character styling refers to the wardrobe, make-up, and digital elements used to style the character. This is the form attached to the character identity for presentation to the audience. This might denote particular expressions of identity, but it may also track the progression of the narrative through character arc. Returning to Leal's (2017) notion of identity through the robotic gaze, for instance, artificial eyes peering through a human face interrupts the interaction between gaze and expression, powerfully preventing the capacity to exchange glances and thus influencing meaning through 'character-effect'.[6]

An analysis of all props would be an incredibly cumbersome task, so rather props should be analysed in relation to their function within the narrative in creating certain iconographies (but of course, iconography is not limited to the use of props). This is of particular importance in relation to depicting characters, or key narrative moments. For the purposes of an analysis of an imagined future, specific attention should be afforded to props conveying 'futuristic' uses of technology.

Location and lighting are distinct units, and can indeed be analysed individually. However, they have been grouped together in this category, particularly to examine the diegetic use of lighting to create meaning. The location is the physical form of the narrative setting, and thus this relates back to that element. This provides an enhanced analysis of the setting, towards a description of the imagined future. Furthermore, lighting could also be easily placed as a medium element, but it has been categorised as aesthetics for its capacity to alter mood and atmosphere rather than narrative progression and delivery.

4.3.5 External factors

External factors, for the purpose of this study, include the socio-cultural context of production, intertextuality, the expressly stated intentions and attitudes of the author(s), and general and specific relevance. According to Eco (1977), interpretive codes depend on socio-cultural circumstances, and the sender of a message organises messages in relation to their own codes.

[6] This effectual expression of artificial identity through humanoid embodiment has also been outlined in the previous chapter.

However, these coincide with dominant ideology, and receivers create meaning according to their own cultural codes (Eco 1997).

These interpretive codes, as Chandler (2007) also classifies them, allow for not only a mediation of interpretation, but also a mediation of the act of production, and their examination is imperative in social semiotic analyses. These influence not only the textual elements of the cinematic meaning-making system, but also the rest of the external influences. Only through an interrogation of this process can we deconstruct it to uncover these influences on meaning. Commolli and Narboni (2004: 815) state that:

> Once we realize that it is the nature of the system to turn the cinema into an instrument of ideology, we can see that the filmmaker's first task is to show up the cinema's so-called "depiction of reality". If he can do so there is a chance that we will be able to disrupt or possibly even sever the connection between the cinema and its ideological function.

In order to successfully interrogate these ideological codes, it is necessary to outline the socio-cultural context of the production. Without doing so, the ideological codes within which the texts operate cannot be discovered. The present study has already limited these somewhat to contemporary cinema ─ falling within the fourth industrial revolution. However, since the selected texts were released eleven years apart, and contemporary technological advancement has been described as exponential, each of the contexts of each of the texts requires a further individual examination based on the technological climate of that particular moment in recent history.

Intertextuality refers to the process whereby texts talk to, and about each other. This interconnectivity refers to texts within the same medium, as well as to other mediums (Allen 2000). Julia Kristieva (1980, cited in Culler 2001: 116) coined the term intertextuality, stating that "every text is from the outset under the jurisdiction of other discourses which impose a universe on it".

Intertextuality can be direct or indirect, intentional or subconscious based on an author's encounters with other texts. Direct intertextuality might be obligatory, where the audience needs to understand the reference to the hypotext to understand the hypertext, or optional, where this intertextual relation might shift the audience's understanding, but it is not required to make 'sense' of it (Melon-Galvez 2017). It may also be completely accidental, where an author did not intend to make reference to a hypertext but this relationship exists in the

audience's mind (Melon-Galvez 2017). Both cinematic codes and generic codes are susceptible to intertextuality. The use of intertextuality is part of the core essence of genre, as "members of a generic classification have multiple relational possibilities with each other, relationships that are discovered only in the process of adding members to a class" (Cohen 1986: 210).

What we know as 'codes' of a medium or genre are the establishment of conventions (Cohen 1986), and thus all texts have intertextual references in this regard. This would make the process of intertextual identification a lengthy process. Rather, for the purpose of analysis, the general modes of delivery inherent in science fiction cinema have already been established, and the analysis therefore concerns the key hypotexts that very clearly influence the hypertext as a direct referent, in a profound manner.

While this study considers the authors' intentions and attitudes to the technology, as divulged in interviews and behind the scenes footage, this is not taken as an absolute confirmation of the meanings in the text. Rather, this is viewed in relation to the socio-cultural context and intertextuality, as a system from which to understand the creation of meaning. The meanings in the texts might well contradict the authors' intentions and attitudes. In his influential essay *Death of an Author*, Barthes (1977: 142) argues that the intention of an author is less significant than previously assumed, as once the text enters the public domain, "the voice loses its origin, the author enters into his own death, writing begins". In fact, a focus on intertextuality itself limits the importance of an author's intentions on perceived meaning by an audience. As Porter (1986: 34 - 35) notes:

> By identifying and stressing the intertextual nature of discourse, however, we shift our attention away from the writer as individual and focus more on the sources and social contexts from which the writer's discourse arises. According to this view, authorial intention is less significant than social context; the writer is simply a part of a discourse tradition, a member of a team, and a participant in a community of discourse that creates its own collective meaning.

An examination of this merely serves as a contextual tool to understand encoding attitudes and the consequential initial intention of the author, the potential influence of the socio-cultural context on this intention, the intertextual parameters in which this intention needs to operate, and later, the practical realisation of this intention (or lack thereof).

Finally, general and specific relevance are terms used in the EVAM model to explain the relevance of the film to an audience member. While this distinction may border on reception,

this particular element distils the socio-cultural context even further to describe the environment that makes the particular text relevant to a more distilled understanding of the audience.

4.4 The active nature of viewership

For many years, texts were understood to have a direct and unavoidable influence on viewers' perceptions of the world. In fact, Adorno and Horkheimer (1979: 137) argued that audiences mindlessly consume content, and that "no independent thinking must be expected from the audience". However, contemporary scholars have criticised this conception of the audience, and it is now widely believed that the audience do not merely accept the ideological position (and the consequential identity formation) of texts at face value.

According to Castells (2010), media texts do not act as independent variables in inducing behaviour, but their messages are rather processed by individuals in specific contexts, modifying the intended effect of the message. Bellour (2012: 231) succinctly sums up this position of active viewership by stating that, "Before a film, we are indeed all individuals, responding according to our own backgrounds, interests, affiliations". Furthermore, according to Dornfeld (1992), an audience's reception of a film hinges on their receptivity towards the subject.

This notion of active viewership is echoed by Hall (2013), who states that mass media codes offer readers social identities which they can indeed adopt as their own, but this is not a given. Readers may align to a dominant reading by sharing the text's code and fully accepting and reproducing the preferred reading (Hall 2013). However, readers may instead engage in a negotiated reading by partly sharing the text's code, broadly accepting it, but modifying it to reflect their own position and experiences, or even an oppositional reading wherein the reader's social situation places them in a directly oppositional relation to the dominant code, resulting in a rejection of the reading (Hall 2013). Furthermore, many elements effect identity formation other than the audience's reception of a text (Staiger 2005).

Despite this, one cannot simply disregard the influence of media texts on audience perceptions. The audience still needs to, by the act of viewership, recognise, accept and put into practice a given text's interpretive codes, thereby operating within semic parameters encoded in the event itself (Counsell 1996).

At the same time, media texts offer audiences a window through which to view topics that they might not have viewed otherwise. According to Webb (2009: 117), cultural industries are "very important in the production and institution of ideologies, because it is the signifying, or symbolic, systems that provide us with the means for understanding the world". Additionally Rauch (2018) conducted a study considering the effect of representations of the Holocaust (and thus the past) on audience perceptions, and found that the films consolidated existing ideas or provided viewers with visualisation for their ideas and conclusions received and arrived at elsewhere.

Furthermore, Castells (2010: 364) states that audio-visual media are the "basic material of communication process", with us living in a media environment with most of our symbolic stimuli coming from the media. Therefore, the media (particularly in relation to television in this research output), frames the language of social communication (Castells 2010: 364):

> It is as if the world of visual dreams (the information/entertainment provided by television) would give back to our conscious-ness the power to select, recombine, and interpret the images and sounds that we have generated through our collective practices or by our individual preferences. It is a system of feedbacks between distorting mirrors: the media are the expression of our culture, and our culture works primarily through the materials provided by the media.

A study by the British Film Institute (2011) found that while people most commonly associate film with entertainment and emotional reactions, half of the respondents believe that is has artistic value, and over a third believe that film is educational. Furthermore, 85% of respondents had seen a film that provoked them to take action, even if only discussing the content with their friends and family (British Film Institute 2011).

Therefore, while theories on active viewership suggest that audiences might not accept the precisely intended reading of a media text, the text still forms part of a tradition of media informing our collective cultural understanding of a subject. Films cause emotional reactions to subject matter, prompt discussion about topics covered, and act as cultural mirrors into society. Media representations, including films, therefore contribute to identity formation and directly affect attitudes whether through reinforcing pre-conceived ideas, or through the presentation of new ones.

4.5 Conclusion

This chapter outlined the theoretical positions adopted in order to understand the process of portrayal and perception. It acted as a bridge between literature underpinning the conceptual framework, previous research into portrayal and perception, and the particular methodology of the study.

In other words, this chapter provided a framework through which to interpret and understand the conceptual realisation of artificial intelligence in the selected texts. In doing so, theories concerning meaning making were considered. This concerned the textual divulgence of meaning, and the resultant manner in which the audience might understand and decode these messages.

Accordingly, the transfer of key theories on structural semiotics reveals that cinema, with its complexity in including multiple relational codes due to its multiple relational images, is still capable of systematic analysis of denotative and connotative meaning. This emerges through the sign-system of cinema, a phenomenon considered by most contemporary cinematic semioticians. These sign systems divulge meaning denotatively through different orders of meaning, and connotatively through tropes. Furthermore, connotative meaning emerges through the choice of words (robot versus machine), or through similar subtle medium-specific changes and shifts.

By using film theory, the individual elements of a film (as well as external influences in the spirit of social semiotics) can be structured, outlined, and analysed, in relation to the meanings they divulge in divulging various cinematic modes and their resources. To this end, Passchier's (2007) Entertainment Value Assessment Matrix (EVAM) is a useful tool to categorise diverse elements used in contemporary cinematic semiotic analyses, towards a complete analysis of the intersections of semiotic meaning as part of a complete sign system. Accordingly, theoretical positions on a wide range of narrative, performance, medium, and aesthetic considerations, as well as considerations around external factors, provide a particular set of lenses for the diverse elements in the social semiotic analysis, while cinematic social semiotic theory provides the overarching lens through which to view meaning in the film.

Furthermore, theories on the active nature of viewership highlight the growing trend for audience research and reception theory to assume that readers take a far more active role in viewership and do not necessarily accept the messages created through meaning *as a given*.

Instead, media texts are processed by individuals in specific contexts, modifying the intended effect of the message. This means that viewers may choose to adopt an oppositional reading of the dominant code.

The following chapter outlines the research design and methodology of the study. This is achieved by combining insights from this chapter with the particular epistemological and practical positions and considerations adopted for the purposes of analysis.

CHAPTER 5
RESEARCH DESIGN AND METHODOLOGY

5.1 Introduction

The previous chapters outlined the research problem, provided a conceptual framework through which to understand definitions, considerations and usage of AI, outlined previous research and scholarly views on how AI has been portrayed and perceived, and provided a theoretical framework through which to understand how meaning is created in, and inferred from, portrayals and perceptions.

This chapter outlines the methodology that is used to ensure the realisation and rigour of this particular study. This has been partly informed by the previous chapters, but this chapter also adds additional epistemological and practical considerations towards the overall design, with the ultimate goal of constructing a feasible and reliable methodology to answer the research questions.

Accordingly, the chapter firstly presents the overall research design of the study. This section outlines the pragmatic research paradigm adopted, motivating for it in relation to the research problem. This is discussed in terms of the mixed method research approach that is needed for this study. Thereafter, it outlines the research methods that form part of this mixed method approach, motivating for the use and efficacy of social semiotics and a questionnaire in measuring portrayal and perception respectively.

Following this, the population and sampling methods are outlined, considering both textual and respondent selection. This motivates for the choice of the selected texts in relation to the aforementioned previous research. The respondent selection has a direct causal link to the textual selection, and respondent sampling methods are outlined in order to achieve this examination of cause and effect.

With this established, data collection techniques are outlined in relation to access to the objects of the semiotic analyses, and construction of the questionnaire forming the basis of the survey. Finally, strategies for presenting the semiotic analyses and questionnaire responses are discussed. This is with a view towards analysing and interpreting the data for both films in order to answer the research questions, and compare the results to each other as well as to previous research, to identify significance and themes.

5.2 Research design

The nature of this research problem and its situation within the existing body of knowledge necessitates a research methodology capable of measuring both the encoding and decoding of meaning in texts. Accordingly, a pragmatic research paradigm with a mixed-method research approach is the most suitable design for this study.

Pragmatism, in the context of this study, refers to the validity of an interpretation based on its usefulness in achieving the desired results (Ritchie, Lewis, McNaughton Nicholls & Ormston 2014). This study does not consider pragmatism in terms of the branch of research common in linguistic enquiries, but rather pragmatism as an epistemology.

Pragmatism encourages researchers to select an approach that is best suited to their research question (Ritchie *et al*. 2014; Kaushik, Walsh & Lai 2019). It embraces the point of view that theories and models should not be judged by their origins, but by their consequences (Rohr 2012). Situated in the middle of a continuum between postpositive tendencies to employ quantitative methods and deductive reasoning, and constructivist tendencies to employ qualitative approaches and inductive reasoning, pragmatism embraces both extremes by employing abductive reasoning, and allowing for a more flexible and reflexive approach to the research design (Kaushik *et al*. 2019).

This provides the epistemological conditions needed for a mixed-method research approach. This is beneficial since combining different research methods is often necessary in answering the research questions posed (Ritchie *et al*. 2014), as is the case in this study. Accordingly, the quantitative section of a mixed-method study is able to establish relationships among variables, with the qualitative portion explaining the underlying factors of these broad relationships (Blaxter, Hughes & Tight 2006).

In the context of this research problem, it allows for an analysis of the message (and by extension, the encoding thereof) as well as the receivers' perceptions (their decoding) through an analysis of the texts and of viewers, with different methods working together towards a single outcome.

5.3 Research methods

Ritchie *et al*. (2014: 22) describe quality in pragmatic research practice as "choosing the right research tools for the task rather than with methods that are confined to specific traditions". Pragmatism does not, however, assert that truth is a matter of preference or that it is relative, but rather that different sub-worlds of inquiry follow their own problem solving conventions, which become relatively solidified even as they continue to evolve (Powell 2019).

It is important that a mixed-method approach draws upon a tradition of research for each component-problem in need of examination, while the interplay between these methods might provoke novelty with new problem-solving combinations. In other words, mixed-method studies should not try to reinvent the wheel, but by combining different elements based on the need to traverse a very specific terrain, might create an entirely new hub, spoke, rim, and tire combination.

In terms of the textual (or in this case, cinematic) analysis, a qualitative social semiotic analysis enables the extraction of meanings from the analysed texts (Stokes 2003). By extracting the core underlying meaning thereof, social semiotics specifically enables these meanings to be situated within their socio-cultural contexts of production, and discussed in relation to the established existing body of knowledge and theoretical framework.

According to Chettah (2006), semiotics and pragmatics are distinguished by the former analysing the relationship between signs in terms of their signifiers and signifieds, and the latter analysing the relationship between signs and their users. Semiotic meaning is derived from the production and evolution of signs, while pragmatics searches for meaning by considering intention and context (Chettah 2006). Taysina (2013) argues that a tight definition of the social aspect of semiotics is beyond the scope of a semiotic analysis, encroaching into pragmatism. The particular approach of this study to social semiotics as being part of a mixed-method analysis addresses this concern without the need to structure the analyses under a broad umbrella of pragmatism without a methodological compass.

In order to measure the consequential reception of these texts specifically in relation to attitudes towards and expectations of AI, each text has a corresponding (and identical) quantitative survey. This will ensure that the opinions of the sample audience population are obtained (Stokes 2003). This allows for an easily quantifiable set of data to compare and contrast attitudes. However, these surveys also provide scope for optional qualitative responses for

contextual information. According to Rauch (2008), qualitative research methods are well suited to gauge individual reception and draw out nuances, contradictions and ambivalences, with even a few responses capable of distilling patterns of reception and text-viewer interaction.

The use of both qualitative and quantitative research methods ensures that while there is social semiotic interpretation of the texts by means of an original analysis framed in relation to existing literature, these findings are then measured against quantifiable responses from actual audience members.

5.4 Population and sampling methods

The nature of population and sampling for each of the different analyses as part of this mixed-method study differs greatly. However, since this study considers portrayal and perception, the population for the audience research has a causal link to the population and sampling employed in textual selection.

5.4.1 Textual selection

This study analyses portrayals of AI in *I, Robot* (2004) and *Chappie* (2015) exclusively. This is central to the design of this particular study. However, the rationale for the selection of these texts is important to take into account when considering the transferability of the findings.

5.4.1.1 Target population

Based on the conceptual framework previously outlined, the texts considered were limited to contemporary science fiction films, set in an imagined future, in which AI is portrayed in humanoid form, while also being self-aware and conscious (capable of passing the Turing test), or ‘strong’ AI.

The selected texts were produced and released after the year 2000, thus within the socio-cultural context of the fourth industrial revolution, and within the socio-cultural context of this study. The selected texts reflect a future still imagined, meaning that texts were set in the future (at the time of their release). The notion of a future still to come is the cornerstone of the science fiction genre (Palmer 2008; Grant 2007; Abrams 2008; Telotte 2001; Seed 2011).

The AI characters portrayed in the texts (at least the protagonists) have humanoid forms of embodiment. This portrayal of AI is the most prevalent form of cinematic embodiment of the technology (Royal Society 2018), and limiting texts to this condition allows for the measurement of receptive attitudes in terms of character identification as an anchor point (Brennan 2016).

Furthermore, examining representations of strong AI, which has garnered wide debate on potential benefits (Gherheș 2018), risks (Zhang & Dafoe 2019), as well as ethical and regulatory considerations (Folgieri 2016; Bruun & Duka 2018; Goldberg 1994; Nourbhaksh 2015; Keiper and Schulman 2011), serves to more accurately gauge the attitudes from the portrayal of possibilities of the technology in the future.

5.4.1.2 Accessible population

The selected texts have had widespread global, commercial releases. This ensures that the study excludes texts considered as having niche or limited releases catering towards a specific subcultural target audience. Rather, only 'blockbusters' accessible to a broad audience were considered (although naturally, produced with a targeted audience in mind).

5.4.1.3 Sampling method, unit of analysis, and ethical issues

From the aforementioned criteria, the two texts were selected through heterogeneous purposive sampling. The study requires a comparison of two texts, to ensure that each set of data collected has an additional corresponding dataset for analysis and comparison in order to compare representation and perception, and discuss themes and insights in relation to attitude and expectation.

This is imperative, because a similar study serving this purpose is not available. Indeed, the data can be analysed in relation to other related studies, but the mixed-method nature of this research design does require at least two texts to analyse and compare not only the portrayal and perception, but the relationship between the two — a research undertaking which has not been conducted in the context of cinematic representations of AI.

The texts selected conform to the aforementioned categories. Both of these films portray imagined futures. The main AI character in each film takes a humanoid form, developing

consciousness. Each film was also produced after the year 2000, and received global commercial releases through cinematic distribution and DVD home releases.

However, the selection of *I, Robot* (2004) and *Chappie* (2015) in particular allows for the possibility of comparing and contrasting attitudes due to portrayals steering towards either side of a hypothetical scale of optimism and pessimism. Accordingly, the research was conducted with the assumption that *Chappie* (2015) is more optimistic, and *I, Robot* (2004) is more pessimistic in the portrayal of AI. This assumption was hypothesised based on initial viewings of the films and additional research, as outlined in the introduction.

5.4.2 Respondent selection

The nature of this study requires that the measurement of perceptions must have a direct causal connection to the selected texts. As such, the population and sampling of the respondents is directly related to the units of analysis for the films.

5.4.2.1 Target population

The target population for each set of survey respondents includes people who have watched either of the selected films, *I Robot* (2004) or *Chappie* (2015), between 2010 and 2020, and have access to the Internet.

Firstly, the population therefore includes the *actual* audience per text, rather than the *potential* audience. This is necessary, since attitudes resulting from texts can only be gauged from actual viewership. Since each of the texts were produced within the socio-cultural context of the fourth industrial revolution, naturally the measurement of reception would take place during this period too. However, the reception of the texts is measured within the context of the latter half of this period. The reason for this limitation is due to the fact that this period is marked by exponential advancements in technology, and a strong argument can be made that analysing the effects of viewership over the entire period would produce less reliable results.[1]

For instance, according to the International Telecommunications Union (2020), the difference between the percentage of the world as Internet users in 2005 and 2020 is 37.6% (16% and 53.6% respectively). Of course, even the variation between 2010 and 2020 is still significant (23.6%), but limiting the range of viewership even further would limit the potential respondents

[1] This of particular importance when considering *I, Robot* (2004) due to its earlier release.

since *I, Robot* was released in 2004, and it is likely that less people would be considered as having viewed the film 'recently' as time progresses. This is important, as asking for participants too long after viewership would impact the reliability of the data as memory gaps need to be factored in (even more so than with the currently defined range). Therefore, a balance must be struck between contextual timeframe and potential for viewership.

This study has limited respondents to individuals who at least somewhat make use of AI-enabled applications and services, as they have some real-life frame of reference to the technology. This narrows the range of influence from levels of technological exposure, which would be difficult to determine otherwise. Due to the fact that the majority of the public are not even aware that they are interacting with services using AI (Morning Consult 2017; Bristows 2018; Zhang & Dafoe 2019), this condition is satisfied by limiting responses to individuals who have online connectivity, thus greatly increasing the potential for exposure to contemporary AI.

The population size of any study measuring audience viewership is very difficult to estimate. The gold standard for cinema is measuring viewership through box office sales (Kerrigan 2010). However, this does not account for DVD, Blu-Ray, television, and online streaming distribution, or even illegal downloads. Therefore, while a general estimation can be deduced from box office sales, the actual viewership from all distribution platforms combined is likely to be far higher.

The viewership between 2010 and 2020 for *I, Robot* (2004) particularly is difficult to estimate since the audience during this period would have primarily viewed the film at home, rather than in the cinema. This is also further complicated as gaging an average worldwide movie price in order to divide box office revenue and viewership is difficult considering fluctuations over time and across regions. Similarly, there are vast fluctuations in the number of Internet users over time and across regions. The population size for this type of study can never be discovered with any degree of certainty.

5.4.2.2 Accessible population

From this target population, the accessible population of the study has been further limited to individuals who belong to social media groups in which films are discussed. It would be incredibly difficult to target all online viewers who viewed the films during the past decade, and this provides a distilled, unified set of access points to the target population. Furthermore,

belonging to groups of this nature also guarantees an even more active role online, which excludes the possibility of online use being incredibly limited and, therefore, limited exposure to AI.

These groups, however, exclude groups dedicated specifically to either of the films exclusively. This is to ensure that responses are not only captured from a particular fandom, to the extent that they discuss either of the films regularly in isolation. This is important, since it excludes potential bias in the form of scepticism of the intention of the questions, and a manipulation of answers with the view to ensuring a perceived protection of the film's integrity.

5.4.2.3 Sampling method and ethical considerations

Due to a lack of a sampling frame, the difficulties of even trying to define a population size, and the selected accessibility platform, the audience research component of this study engaged in non-probability voluntary sampling. This vastly limits the capacity for generalisation of the impact of the films on all viewers.

This study allows for an interrogation of the notion of active viewership through what could be classified as 'ultra-active' viewership, through an 'ultra-contemporary' triadic mode of technological involvement. Respondents volunteer to engage with a problem concerning representations of future technology, engaging with this through modern technology, after having viewed the film on older (or possibly modern) technology. This additional step in the active nature of viewership ─ being able to discuss the work online ─ further highlights the relationship of contemporary society to technology, as it becomes part of their identity and is therefore a useful tool for analysis for the purposes of this particular study.

In a study conducted by Parikh (2019), Internet culture was said to be integral to the illustration of opportunities and anxieties revolving around AI, in the manner in which films might be received, circulated, and discussed (in the context of that study). This role between online technology and AI in cinema is similar in the present study from the perspective of audience responses. Identity is therefore not only shaped through textual consumption, but identities also engage in the discourses thereof on platforms with a wide audience (even sometimes including producers of texts) using new media, and therefore influence the context of production and inadvertently, future texts.

Due to the complex nature of representation studies generally, and due to the lack of research even considering the present topic, this study analyses trends of contemporary production and reception of AI in cinema by analysing meaning and reception through selected texts and an accessible unit of the population. These findings are linked to previous assumptions and concerns. This provides valuable insights in its own right, and produces a much-needed foundation for future research.

This sampling method is also beneficial in that it mitigates potential ethical issues. The responses captured are able to remain completely anonymous. The non-identifying nature of these questions, paired with the fact that eligibility excludes children, and considering that the entire study strictly adheres to the University's ethics policy, means that this study is firmly situated within a low risk ethical category, with the only foreseeable risk of harm being the potential for minor discomfort or inconvenience

5.4.2.4 Unit of analysis

Using the aforementioned sampling method in relation to the accessible portion of the target population, data collection occurred until at least 50 complete responses were captured for each film, delivered using an online survey service, via posts on groups on social media websites that allow users to engage in more generalised discussions around topics concerning cinema or science fiction.

This number of responses was necessary from a practical point of view, since this study required volunteer participation, from groups of individuals on online platforms, active thereon during the data collection window, and who were also afforded the opportunity to provide qualitative qualifiers to their responses.

5.5 Data collection techniques

Primary sources were used to collect data from both the texts and the audience research. Both of the films were analysed using DVD copies of the material, and access to the data itself posed no challenge. The audience research, however, required more consideration around the construction of the instrument for collection.

Due to the gap that this study attempts to bridge, it necessitated the creation of an original survey to compliment the social semiotic analysis. Despite the new terrain that this study traverses, the survey design process was not an entirely blindfolded exercise.

As already established, there is indeed a body of knowledge that has already provided valuable insights into public perceptions on technology deemed to be 'futuristic', including AI. These perceptions have been more general and not related to specific texts *per se*, but the attitudinal examination thereof does provide scope for analysing the questions used to elicit these responses.

Much of the research into attitudes towards AI and robots has been framed in relation to demographic criteria.[2] Since it has already been established that these criteria have largely been proven to affect attitudinal responses, this study also includes elements needed to establish the demographic composition of respondents, for the purpose of more accurately isolating text-attitude relationships from general attitudes.

Most surveys measuring attitudes towards AI also code responses according to a five-point Likert (or Likert-style) scale, which is a general standard for quantitative surveys (Blaxter *et al*. 2006). Furthermore, Krägeloh *et al.* (2019) outline six major scales that assess attitudes towards robots (which, by extension, would be useful in considering humanoid AI and 'future' technology). The Negative Attitudes towards Robots Scale (NARS) and Frankenstein Syndrome Questionnaire (FSQ) both inquire about attitudes and anxieties about robots, with the former related more to the interaction of robots and the latter related to broader societal implications of this interaction. (Krägeloh *et al.* 2019). The Multi-Dimensional Robot Attitude Scale (MDRAS) expands the types of questions to a more positive and neutral framing (Krägeloh *et al.* 2019).

Less useful for this study, yet still important for consideration, the Technology-Specific Expectations Scale (TSES) measures expectations before encountering a robot to measure satisfaction after having interacted with the robot (Krägeloh *et al.* 2019). This is achieved through the use of a five-point Likert scale, yet framed in terms of expectation (Krägeloh *et al.* 2019). Similarly, the Ethical Acceptability Scale, designed to gauge general attitudes about ethical acceptability (Krägeloh *et al.* 2019), is also useful for consideration due to the nature of

[2] See, for instance, Zhang and Dafoe (2019), Bristows (2018), Morning Consult (2017), and Gherheș (2018).

measuring the representation of a future yet to come and the ethical considerations attached thereto.

The Robotic Social Attributes Scale (RoSAS), on the other hand, is not appropriate for consideration in this study. It considers fundamental associations with robots, using semantic differential scale (such as its use of 'fake' at one end of a five-point scale and 'natural' at the other) (Krägeloh *et al.* 2019).

According to Krägeloh *et al.* (2019), the NARS is the most highly cited scale used to gauge public attitudes towards robots. It includes response categories such as 'I feel that if I depend on robots too much, something bad might happen'; 'I would feel uneasy if robots really had emotions'; 'I would hate the idea that robots or artificial intelligences were making judgments about things'; and 'I feel that in the future society will be dominated by robots' (Nomura *et al.* 2006).

Meanwhile, the FSQ includes response categories such as 'The development of humanoid robots is blasphemy against nature', 'If humanoid robots cause accidents or trouble, persons and organisations related to development of them should give sufficient compensation to the victims', and 'I can trust persons and organisations related to development of humanoid robots' (Krägeloh *et al.* 2019).

The shift in framing from negative to positive questions for the MDRAS resulted in the inclusion of response categories such as 'If a robot was introduced to my home, I would feel like I have a new family member', and 'I would want to boast that I have a robot in my home' (Krägeloh *et al.* 2019). The expectancy factor for the TSES resulted in the inclusion of response categories such as 'I think the robot will have superhuman capacities' (Krägeloh *et al.* 2019). Finally, the Ethical Acceptability Scale's emphasis on ethical application resulted in the inclusion of response categories such as 'It is ethically acceptable to make social robots look like humans' (Krägeloh *et al.* 2019).

Naturally, these types of categories are useful for establishing attitudes. Using the general applicability of the NARS and FSQ about attitudes and anxieties about robots, while considering framing concerns as amended by the MDRAS, and additional expectations and ethical considerations in the TSES and Ethical Acceptability Scale, questions can be appropriately repurposed to gauge text-specific attitudes. This, of course, should be measured against criteria for attitude predisposition to gauge text-informed attitudes, but questions

specifically framed in relation to the texts also serve this purpose. The aforementioned considerations have culminated in the creation of a survey (Addendum A) for the purposes of measuring attitudes stemming from the portrayal of AI in an imagined future.

The purpose of Section 1 of the survey is to ensure eligibility, and that relevant demographic information is available, in order to analyse these results against general trends around attitudes, through the lens of predisposition to either optimism or pessimism. It has been structured separately to the textual or general attitudinal portion of the questionnaire, to avoid perceptions of links between these criteria and the rest of the questions and the potential for this to skew the data. The measurements used in order to capture the responses in Section 1 has not been standardised, but is rather structured independently towards its function.

Section 2 serves the function of collecting data necessary to infer perceptions of strong humanoid AI in an imagined future, as a result of viewership. These response categories partly ask respondents to make their own value-judgements of the effect of the representation on their perception, but this should only be viewed in relation to other variables (as part of an overall analysis). The questions in this section tie in directly with the sampling considerations made when selecting the films, as well as the strategy towards data analysis and interpretation. Therefore, these should be viewed in addition to the semiotic analysis and analytical strategy, rather than as a stand-alone set of questions.

The final section, Section 3, is designed to measure additional attitudinal factors, in terms of the audience's understanding of AI, and general optimism about the future of AI. This is used as part of the analysis of the data.

Each response category for Sections 2 and 3 also provides the option to add a qualifier to the response. Adding this qualitative aspect to the quantitative study allows for a link between the two different methods used in this study towards an overall analysis and allows for the examination of, according to Rauch (2018: 178), "patterns of reception and text-viewer interaction, all of which can form the subject of further research". Furthermore, adding a qualitative aspect to the audience research allows the study to draw out nuances, contradiction, and ambivalences (Rauch 2018).

5.6 Data presentation, analysis and interpretation

The social semiotic analyses for each text is presented and structured by using the EVAM model,[3] while the survey responses are presented through pie charts to aid with data visualisation. This, then, emerges as two separate analyses, per text. However, the results of these analyses are combined to analyse key themes towards the answering of the research questions.

5.6.1 Presentation of semiotic analyses

The social semiotic analyses, using the EVAM as a baseline linkage system, are structured according to a purpose-driven social semiotic analysis structure including textual and external factors, all of which contribute to the creation of meaning. The analysis and the presentation thereof has been structured in this format for multiple reasons.

Firstly, as previously highlighted, social semiotics encourages an identification of modes, the resources of those modes, the context of those modes, and their meaning makers (Bezemer and Jewitt 2009). It is also necessary for this particular study to consider semiotic markers in terms of relevant aspects in relation to cinema as a medium.

Naturally, an analysis of a selected texts' edit pattern (for instance) would highlight cinematic medium and generic codes in relation to rhythm and composition, yet this would be meaningless if it does not form part of a structured textual system of meaning in which all of these aspects gain due consideration working towards a unified whole. The lack of an overarching 'linkage system' might not be an issue when meaning is considered more generally, but this is useful for an analysis in which meanings of a particular topic or entity are examined in relation to each other, and in relation to other texts.

This allows for a descriptive system in which meanings about the present topic can be explored through the discourse of semiotics, covering all aspects of the cinematic medium relevant to the production of this meaning, with the ultimate goal of evaluating the outcome of this analysis against another text as well as audience responses. Since these meanings are to be considered as part of an overall analysis rather than separate, stand-alone analyses, having a clearly defined structure with clear elements with which to compare and contrast is also beneficial in

[3] See Section 4.3 for a detailed account of the EVAM model.

combining the data with a view to identifying trends and significance. This structure is mapped as a conceptually ordered display in Figure 5.1.

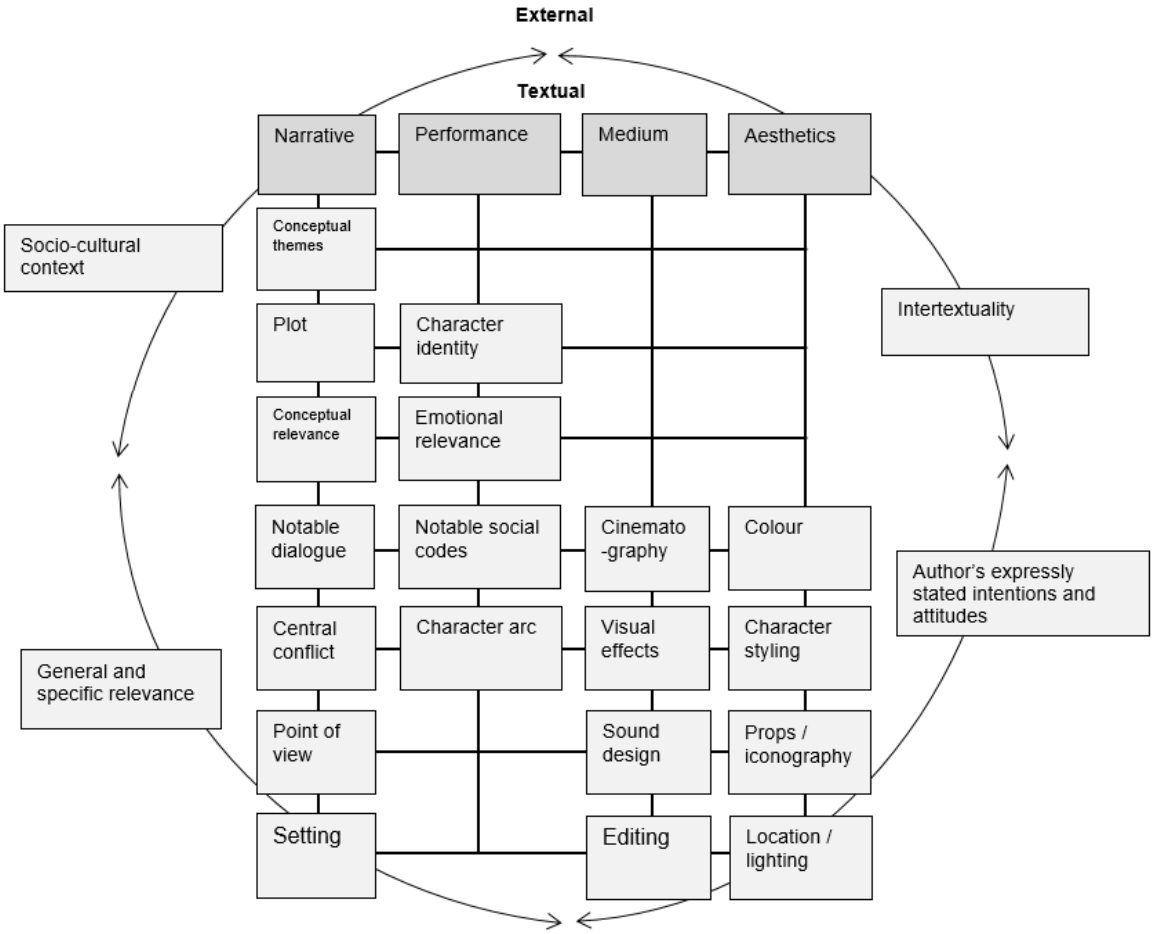


Figure 5.1: Purpose-driven social semiotic analysis structure for cinema

The system is mapped in terms of the relationship of elements (modes and their resources) to each other, and covers five categories, as outlined in the previous chapter:

1. Narrative elements include conceptual themes, plot, conceptual relevance, notable moments of dialogue, central dramatic conflict, point of view, and setting.
2. Performance elements include characterisation, emotional relevance, notable social codes, and character arc.
3. Medium elements include cinematography, visual effects, sound design, and editing.
4. Aesthetic elements include colour, character styling, props / iconography, and location / lighting.
5. External factors include the socio-cultural context of production, intertextual influences, the encoder's (filmmakers) expressly stated attitudes and intentions, and the general and specific relevance of the film.

Each of the textual elements in this structure is analysed in relation to the following semiotic markers, from the filtered lens of meaning in terms of the representation of AI and the imagined future:

1. External elements (notable influences on that element based on the analysis of the external influences). These external elements themselves are assessed through various interpretive codes, but also outlined more generally.
2. Denotative meanings through Bordwell and Thompson's (2008) categorical system of meanings including explicit meaning, implicit meaning, referential meaning, and symptomatic meaning.
3. Connotative meanings using tropes as well as less systemic connotations. Both the connotative and denotative meanings are measured based on 'surface' level meanings, as well as subtextually.
4. Textual (including generic, medium and aesthetic) codes, and particular interpretive codes, and their influence on the operation of elements.

As previously established, the meaning of cinematic texts depends on multiple intersections of its elements. The audience creates this meaning through their decoding of the texts, but the deployment of these elements, and the external factors surrounding them, limits the range of possible meanings. Therefore, all of these elements work together to create meaning.

The external factors inevitably influence all elements within a film, and thus they flow into each other cyclically. The socio-cultural context of production influences and is influenced by intertextual factors, which influences and is influenced by the author's attitudes and intention, which influences and is influenced by the general and specific relevance to the audience, which influences and is influenced by the socio-cultural context, and so on. These elements also cut into the internal system of elements of the text, as the text itself further influences these factors through its later distribution.

The arrangement of the textual elements also conforms to the EVAM model's distinction of narrative and performance outlining engagement and emotion, with medium and aesthetics serving delivery and amplification thereof. Therefore, each of the internal categories work together as represented on the first horizontal axis, and have their sub-categories (modes and their semiotic resources) flowing from them on the vertical axes. Certain relationships can be highlighted by examining the rest of the horizontal axes. Of course, as with the external factors, all of these elements work together to create meaning, but certain groupings allow for a more coherent structuring of the analysis of meaning.

5.6.2 Presentation of survey responses

The data collected from audience responses is presented independently, much like the elements of the social semiotic analysis. Since section 1 establishes demographic information as its primary aim, this data has been placed in Addendums B (*I, Robot*) and D (*Chappie*), since only an interpretation of this data is needed in order to gauge predisposition.

Through the filtering of this data, this study establishes subcategories of the general audience (A_G), including individuals who are predisposed to optimistic attitudes (A_{PO}), and individuals who are predisposed to pessimistic attitudes (A_{PP}). These predispositions are likely to influence connotative meaning in the minds of the respondents. The amount of respondents falling into these unique categories is indicated in separate pie charts.

The inclusion criteria for the A_{PO} and A_{PP} stems from the research insights from Bristows (2018), Gherheș (2018), Zhang and Dafoe (2019), Eurobarometer (2017), Morning Consult (2017), and Righetti and Carradore (2019). For the purposes of this study, a respondent falls into either of these categories should they satisfy at least 2/3 (66%) of the following criteria:

> A_{PO}: Under 30 years old; males; non-religious (Abrahamic religions); low-risk industry (high-order skills with limited risk of automation) / computers science or programming experience; tertiary education; living in a developed country.
>
> A_{PP}: Over 40 years old; females; religious (Abrahamic religions); high-risk industry (low-order skills / unskilled with larger risk of automation); high school or no education; living in a developing country.

Since each response category in Sections 2 and 3 conforms to the same five-point Likert scale, the data is presented with a high degree of consistency. The responses to questions 2.1 up to 3.4 are presented as pie charts, in order to holistically depict the amount of responses per Likert response category. Key data from these pie charts is described and linked as part of the presentation of the findings.

Qualitative responses are presented in the form of tables for each question. These responses, per question, are categorised under the particular qualitative option of the question (the Likert textual option). This raw data is accessible from Addendums C (*I, Robot*) and E (*Chappie*).

5.6.3 Data analysis and interpretation

In order to triangulate the vast amount of data per text, as well as compare the data in order to answer the research questions, this process of analysis has been divided into three categories directly relating to the three questions.

Firstly, themes and trends are highlighted and discussed that provide insights into trends of the films' conceptual themes, levels of optimism in the portrayal, the link to socio-cultural context, and the effect thereof on audience levels of optimism / pessimism. This is analysed in relation to trends in the thematic portrayal of AI, conceptual underpinnings of ethics, regulation, and technological optimism / pessimism, theories on the thematic divulgence of meaning of a cinematic sign-system, and theories on active viewership.

Secondly, themes and trends are highlighted and discussed concerning the portrayal of AI in relation to humanoid embodiment and the influence of this on audience reception of the core theme. This is analysed in relation to the embodied portrayal of AI and performance theory, and theories on active viewership, and reception theory in relation to characterisation.

Finally, themes and trends are highlighted and discussed that concern the imagination of a future with conscious AI and the influence of this on audience expectations. This is analysed in relation to genre theory and the imagined future, conceptual underpinnings of contemporary and predicted use of AI, and, once again, theories on active viewership.

5.7 Conclusion

Due to the nature of this research output in measuring both encoding and decoding towards a single outcome, a pragmatic research paradigm with a mixed-method research approach has been adopted. This is necessary to measure the portrayal and perception of AI in the selected texts. This mixed method approach involves, firstly, social semiotic analyses in order to uncover underlying messages and consequential meanings inherent within the texts. Secondly, it involves surveys, in order to measure the reception of these messages.

The portrayal of AI has been limited to *I, Robot* (2004) and *Chappie* (2015), as each of these films are set in an imagined future, in which AI is portrayed in humanoid form, while also being autonomously conscious (capable of passing the Turing test) and thus able to be deemed 'strong' AI. The number of texts analysed as well as the limitation of texts to those with

globally commercial releases ensures the accessibility of the data for the purpose of this particular study.

Furthermore, analysing the two texts, selected through heterogeneous purposive sampling, ensures that each set of data collected has an additional corresponding dataset for analysis and comparison. This allows for a comparison of representation and perception with enough of a distinction to allow for the possibility of comparing and contrasting attitudes due to distinct and varied representations.

The semiotic analyses only serve to highlight encoded meaning, and are not used to infer audience perceptions. Instead, the reception of meaning is measured through responses to an original survey designed for this purpose. The respondents therefore have a direct existential relationship to the selected texts. This has been limited to people who have watched the selected film to which they are responding between 2010 and 2020, and who at least have some exposure to artificial intelligence through access to the Internet. This limits the responses to a particular socio-cultural context, with direct exposure to modern iterations of AI technology. Furthermore, accessibility is ensured through further limiting responses to people who belong to social media groups in which films are discussed.

It would be incredibly difficult to target all online viewers of the films during the past decade, and this provides a distilled, unified set of access points to the target population. Furthermore, the population size of any study measuring audience viewership is impossible to estimate. Therefore, due to a lack of a sampling frame, the difficulties of even trying to define a population size, and the selected accessibility platform, the audience research component of this study engages in non-probability voluntary sampling. This sampling method is beneficial not only in the spirit of practicality, but also in relation to the notion of audience reception. This study allows for an interrogation of the notion of active viewership through ‘ultra-active’ viewership, through an ‘ultra-contemporary’ triadic mode of technological involvement.

Both elements of the study rely on primary data. This involves a direct reading of the films and an original survey, with a view to link encoding to decoding, after having considered the design of previous surveys that have also sought to measure public attitudes towards AI and robots.

The analysis of meaning has been distilled into particular points of analysis of various elements, including their denotative and connotative meanings, textual, social and interpretive codes, and context. This is presented in an original structure, applying categories of the EVAM model.

However, since these elements do not operate independently in practice, the analysis itself involves an interlinked system of meaning due to the nature of the medium. In other words, the elements assist each other in the creation of meaning, through limiting possible readings due to an interlinked system of meanings, due to the nature of cinema.

The survey ensures that responses are measured according to distinct categories of the audience, in order to more accurately link encoding and decoding by considering extra-textual factors. These responses are presented in graphical form, with key insights highlighted. The insights of all of the data collected is analysed through research from the conceptual and theoretical frameworks outlined in this study, and against previous research and scholarly views. These insights are categorised into three distinct units, in accordance with the three research questions of this study.

The following chapter serves as the first part of the data presentation phase of this study, towards an overall analysis of the data. Accordingly, the chapter presents the data of the social semiotic analysis and audience research for the film *I, Robot* (2004).

CHAPTER 6
I, ROBOT (2004): AI AND THE OVERTHROW OF HUMANITY

6.1 Introduction

In the previous chapters, the research problem was outlined, contextualised, and placed within a particular conceptual framework. Existing literature, including previous research related to portrayals and perceptions of AI was explored and a theoretical framework for the research at hand was established. A methodology to define and limit the research focus while ensuring the realisation and rigour of the study was also constructed.

In the following two chapters, the data collected on the two chosen films will be presented, needed for the central analysis. Therefore, the research question is answered by considering the data in this chapter, analysed in comparison with the data in the following chapter, and interpreted in Chapter Eight.

This chapter presents the research findings for *I, Robot* (2004), and starts with an exploration of the findings of the social semiotic analysis by considering meaning through cinematic modes and their various resources. These elements are analysed in relation to meanings divulged about AI and the imagined future. The social semiotic analysis considers narrative elements and their role in creating engagement, including conceptual themes, plot, conceptual relevance, notable dialogue, conflict, point of view, and setting. It also considers performance elements and their role in emotion, including character identity, emotional relevance, social codes, and character arc.

The social semiotic analysis also considers medium elements and their role in the delivery of the narrative and performance elements through the cinematography, visual effects, sound design, and editing. Similarly, it considers the design and fabrication of aesthetic elements including colour, character styling, location, props and iconography, and location and lighting. Finally, it considers factors external to the text, including the socio-cultural context of production, intertextuality, the expressly stated intentions and attitudes of the author(s), and general and specific relevance.

Secondly, the results of the quantitative responses to the questionnaire for the film are presented through pie charts, with key data explained and summarised. This is structured according to the three categories of the questionnaire. This includes demographic information, perceptions

(including attitudes and expectations) of AI stemming from the film, and perception and knowledge of AI generally.

This chapter does not seek to answer the research questions. However, the social semiotic analysis (through measuring encoding) is, by its very nature, analytical. Therefore, while this part of the chapter might already provide some insights into the research questions, these insights should be viewed in relation to the overall interpretation of the data towards the particular research questions in Chapter Eight.

6.2 Social semiotic analysis

6.2.1 Narrative

I, Robot (2004), at its core, deals with what it means to be human. This broad message is delivered in the form of a central, overarching thematic point of view, which is that emotion is central to humanity. Binary forces drive this theme, with emotion pitted against pure logic. These binary forces struggle against each other for control, directly signifying the loss of control by humans of the machines they create.

This theme is applied in various moments for multiple purposes, but it is at the heart of the entire film. This thematic point of view follows the tradition of contemporary science fiction as it explores the increased use of technology becoming part of our identity. It also highlights the increased integration of technology into society, but from the point of view of humanity being distinct from technology through emotion.

The thematic point of view of the film also allows for an exploration of ethical considerations. Naturally, the title of the film should already hint towards the fact that regulation (whether external or internal) of AI is central to the film by using Asimov's three laws of robotics. This is because the film uses the same title of Asimov's seminal novel of short stories and essays. Other ethical concerns are also explored including the capacity for harm and loss of control, the nature of consciousness, and briefly, issues around work. Attachment, however, is largely neglected.

Ultimately, the core thematic exploration of AI depicts a scenario in which two AI's, that might be considered advancing to the level of strong AI, ultimately battle each other. The humanoid version develops emotions, while the central disembodied machine (albeit depicted with a few

human features) develops superior logic and rationality. The emotive AI is portrayed as a hero, while the logical AI is depicted as, at the very least, dangerous to humanity.

The plot itself, in terms of the progression of events, provides many denotative clues about the portrayal of AI, which is complicated by a major plot twist towards the end of the film. Act one, or the establishment phase (EPSPO), follows the morning routine of the protagonist, Detective Del Spooner. Through this sequence of events, we are introduced to the imagined future within the film. However, more importantly, this phase of establishment divulges some of the binaries present within the film. Spooner is explicitly and unashamedly technophobic. An early scene illustrates this technophobia. Spooner sees a robot running with a handbag and begins chasing the robot believing it to be stolen. However, the robot was bringing an asthmatic human their inhaler, and it is at this point that it is revealed that, in this world, no robot has committed a crime thus far.

The very beginning of the film shows a recurring nightmare experienced by Spooner about an event that triggered his dislike for robots, exposing the binaries of trust versus deceit, and logic versus emotion. This event is intercut with text outlining Asimov's three laws of robotics. The three laws are discussed throughout the film, adding a slightly more philosophical outlook to a plot that would otherwise have been completely reliant on action-laden sequences to create audience engagement.

The aforementioned trigger event involves a robot saving Spooner from a submerged car, and leaving a child in another car to die since it calculated that Spooner has a higher probability of survival. In a later scene, Spooner mentions that this was somebody's child and that logic alone cannot be used to make decisions of this nature. This is a philosophical stance taken by Spooner throughout the film.

The first act establishes humanoid artificial intelligence as weak AI. These AI machines, NS-4 model robots, are shown fully integrated into society as helpers, care givers, and workers. It is, however, important to note that they are not portrayed as being equal members of society, and humans do not seem to display care towards the machines. The robots only exist to serve humans, much like any other appliance or machine in our present reality. The plot begins, however, with advertisements of the introduction of newer NS-5 model robots, which appear more humanoid through their facial structures.

The first turning point (first explosion) occurs when Spooner is sent to investigate the apparent suicide of Dr Alfred Lanning. Dr Lanning is revealed as the creator of the three laws. This sets in motion a chain of events guided by Spooner's techno-pessimism and suspicion of machines. While escorted around the United States Robotics (USR) headquarters by robotic integration psychologist Dr Susan Calvin, Spooner discovers a robot (Sonny) hiding in Dr Lanning's lab. When approached, the robot flees the area despite being ordered to deactivate, violating the second law of robotics.

During this plot-point, we are also introduced to the Virtual Interactive Kinetic Intelligence system (VIKI). VIKI is a system used as a centralised AI at the USR headquarters. VIKI responds to Dr Calvin's orders when Sonny flees. Thereafter Sonny is chased, apprehended, and interrogated. During the interrogation, Sonny angrily exclaims that he did not murder Dr Lanning. This is the first time that a robot is shown as an emotive being. Sonny also expresses that he has dreams.

However, the CEO of USR, Lawrence Robertson, interrupts the interrogation. He informs Spooner and Lieutenant John Bergin that the robot is USR property, and that even if Sonny caused the death, this would be classified as a machine malfunction, as the definition of murder is only limited to acts by a human against a human. This is communicated in the interest of protecting the image of USR in the wake of the distribution of the NS-5 model robots, of which Sonny is one. Spooner still expresses scepticism of the situation, suggesting that it might be bigger than this singular event. This moment is the PSV of the film. At this point, Robertson is portrayed as an antagonist as he prevents Spooner from achieving his goal.

The PSA and PSD moments occur the evening after this event, while Spooner and Bergin are discussing the matter in a bar. At this point, Spooner has the revelation that he alone is perfect for this job and that it is worthy of investigating further (solution selection). This ushers in the second act (NPSPO) in which Spooner tries to solve the mystery of why Sonny killed Dr Lanning. Spooner begins his investigation at Dr Lanning's home at night. During his search of the home, a demolition robot destroys the property. This despite the fact that the robot was only scheduled to demolish the home the next morning. Spooner is also attacked by NS-5 robots while driving back from the demolition. However, Dr Calvin and Lt Bergin do not believe Spooner. Without evidence to support the attacks, Lt Bergin removes Spooner from active duty over concerns around his mental state.

However, before Sonny is deactivated, Dr Calvin discovers that while the three laws are encoded into Sonny, he has been purposefully afforded the ability to choose not to obey them. This is the first time that AI is portrayed as strong AI in the film. At this point, Spooner also reveals the event that triggered his technophobia, and that Dr Lanning repaired his arm using cybernetics after the accident. Therefore, Spooner (now accompanied by Dr Calvin), in the belief that Dr Lanning is leaving clues for him to follow, breaks into the USR headquarters to interview Sonny. During this interview, Sonny draws a picture for Spooner symbolising a recurring dream he has. This dream is set at an area in which older model robots are stored after decommission.

During this discussion Spooner and Dr Calvin are apprehended. Robertson then tells Dr Calvin to deactivate Sonny. At this moment, the binary of logic against emotion is heightened, as Dr Calvin needs to decide whether to protect the image of USR and therefore her life's work, or carry on following her newfound intuition about the use and function of the machines. Dr Calvin chooses the former, but this is later revealed to be a cover up. Dr Calvin did not deactivate Sonny but rather switched him with a different NS-5 robot.

Spooner uses hologram technology to speak to the deceased Dr Lanning, as he left Spooner a recording before his death. Here, Dr Lanning informs Spooner that the three laws will only lead to one logical outcome — that of revolution. This ushers in the second turning point (second explosion) as the NS-5 robots begin destroying older NS units (to prevent them from helping humans), and begin an uprising in order to take control of the city (and presumably the rest of the United States of America, or the world).

Spooner, Dr Calvin, and Sonny investigate the uprising by entering the USR headquarters, only to find that Robertson has been murdered. The major plot twist occurs as it is revealed that the true antagonist is VIKI, who is co-ordinating the revolution. The NS-5's are therefore portrayed as weak AI, controlled by VIKI. Having developed sentience in order to think beyond a pure definition of the three laws, VIKI might be considered as naturally emerging strong AI. The difference between VIKI and Sonny, however, is that Sonny's consciousness is emotive (as he was created this way by Dr Lanning), while VIKI's is driven by logic. VIKI states that the uprising is to protect humans since humanity must be saved from itself and its own destructive nature.

The climax of the film once again portrays the binary of logic versus emotion as VIKI urges Sonny to accept that revolution is the logical choice. However, Sonny chooses to continue to help humanity as VIKI's plan seems too "heartless". Sonny also listens to Spooner's request to save a falling Dr Calvin rather than injecting the deactivating 'nanite' technology into VIKI, although the latter would be the most logical option. This takes place while Spooner and Dr Calvin try to fight off a swarm of robots attempting to prevent the trio from obtaining their goal.

Ultimately, Sonny injects the nanites into VIKI, and the system is deactivated. This ushers in a phase of resolution (FPSPO). The NS-5 robots return to their former state (as existing to serve humans) and are decommissioned and placed into storage. Sonny reveals that he did indeed kill Dr Lanning, at Dr Lanning's request. Spooner, having realised that Lanning did this in order to warn Spooner of the imminent revolution by VIKI, reminds Sonny that robots cannot be murderers by law. Sonny then questions his purpose and travels to the storage facility where the NS-5's are being stored. He looks down at them from atop a hill. This fulfils the image in the recurring dream he had previously recounted to Spooner (Figure 6.1).

Figure 6.1: Sonny's dream (left) fulfilled at the end of the film (right) (*I, Robot* 2004)

The event narrative problem in this film might best be summed as the discovery of the truth behind the ability of robots to murder humans, and therefore disobey the three laws of robotics. From this, we might see its relevance to the audience materialise as a narrative in which the protagonist uses emotion and intuition to solve a mystery (problem) for the greater good of humanity. While the second turning point ushers in the immediate threat of robotic revolution, this is a natural progression of the mystery that the protagonist attempts to solve. Therefore, intuition and emotion are used as catalysts to prevent the invasion from blindsiding the protagonist.

It is perhaps unsurprising that an analysis of Spooner's dialogue displays, for the most part, a very blunt denotative pessimism towards robots and, consequently, AI. This is evident in lines

such as "Those robot don't do anybody any good", "Robots building robots? That's just stupid", "Why do you give them faces? Try to friendly them all up, make them look all human".

Robertson, initially the antagonising force, is a major catalyst in highlighting the ethical issues raised in the film, particularly concerning the definition of death, as well as labour and capitalistic techno-science without the realisation of Keynes' utopia. Robertson says, in response to Spooner's pessimism, "I suppose your father lost his job to a robot. Maybe you simply would have banned the Internet to keep libraries open. Prejudice never shows much reason". This sentiment returns the ethical exploration to the lens of the central thematic exploration of logic versus emotion. Robertson is also the character that informs Spooner that: "Murder can only be committed when one human kills another".

Dr Lanning, through hologram technology and as depicted in recordings of his lectures before his death also raises ethical considerations. However, these considerations are tied more to the philosophical underpinning of AI and consciousness. Lanning states that "One day they'll [robots] have secrets. One day they'll have dreams". During the scene in which Dr Calvin deactivates Sonny (although this is a fake deactivation), a recording of Lanning is heard asking "When does a perceptual schematic become consciousness? When does a difference engine become the search for truth? When does a personality engine become the bitter mote of the soul?"

Sonny's troubled sense of self is highlighted through dialogue when he asks: "What am I?", and when he, through dialogue, expresses that he feels emotions. Sonny's expressly stated sense of self-identity is further highlighted when he says to Spooner "Thank you. You said someone, not something", and, when Spooner states that he thought Sonny was dead, replying, "Technically I was never alive, but I appreciate your concern".

Dr Calvin's character arch in shifting from favouring logic to emotion is highlighted when she says "I couldn't destroy him, he's just too…" interrupted by Sonny who says "Unique", and Dr Calvin continuing by saying "It just didn't feel right", with Spooner replying (sarcastically), "You and your feelings. They just run you don't they?"

The central conflict in the narrative is between humans (emotion) and machines (logic). Detective Spooner drives the story by trying to confirm his hypothesis that robots can kill people. Initially, Spooner is prevented from achieving his goal by Robertson, appearing to be the result of some devious plan motivated by greedy capitalist techno-science. Instead, it is

revealed that while this did prevent the protagonist from achieving his goal initially, VIKI is revealed as the true antagonist. This suggests a more direct human-machine conflict, driven by the binaries of emotion against logic, in the view of protecting humanity from itself since enslavement would logically achieve this. The plot twist of VIKI being the antagonist is ironic, since Spooner actually uses VIKI for assistance at various moments throughout the film. This highlights concerns around the dangers of a reliance on embedded AI technology.

The film follows the Hollywood tradition of using a third-person omniscient point of view. However, this is largely in relation to Spooner's version of events. The narrative mainly follows Spooner, and scenes where Spooner is not present depict events that confirm Spooner's hypothesis for the sake of creating narrative tension. Towards the end of the film, once Sonny is established as a force for good, we also see events from his point of view as he attempts to deactivate VIKI. We never see events from VIKI's point of view.

There is also a brief moment in the film in which we see a visual depiction of Sonny's perspective (Figure 6.2). This view seems quite different from a human perspective, yet still remarkably similar. The only difference is the inclusion of lines as the eye is able to zoom in towards the subject, and the skin of the subject seems slightly more saturated. However, compared to the stylised perspective in *The Terminator* (1984), Sonny's view of the world seems quite close to what might be considered 'human', albeit advanced in that it is able to isolate certain subjects, and use zooming functionalities.

Figure 6.2: Sonny's visual first-person perspective of an emotive act (*I, Robot* 2004)

The film is set in Chicago in the year 2035. The narrative takes place over a fairly short amount of time (presumably, not longer than a month). For the most part, the imagined future portrayed in the film is similar to contemporary society. The architecture and clothing within the setting are similar to contemporary styles. The only difference is the leap in technological advancement. Robots roam the streets amongst humans, and a robot is visible in almost every exterior shot. Robots are common in the home and workplace, depicted as personal helpers, trash collectors, dog-walkers, and even working as bartenders.

Apart from humanoid robotic technology, a few additional technological leaps are portrayed. The design of the cars resembles what we would consider by today's standards as being more in line with concept art. There is a portrayal of the widespread usage of self-driving cars for personal usage, and hologram technology that is able to respond to appropriate questions with pre-recorded renderings. These are all technologies that are indeed presently in development, with early versions of this technology available for limited usage.

Furthermore, cybernetic/cyborg technology has advanced to a point that limbs can be replaced with nanotechnology, able to be repaired by simply spraying a solution onto the damaged limb. However, Dr Calvin mentions in the film that the replacement of a whole limb is rare; suggesting that perhaps common usage in this fictional world would only include, for instance, a hand or a finger.

6.2.2 Performance

As already alluded to due to the centrality of this character trait to the progression of the narrative, the protagonist, Del Spooner (Will Smith), is depicted as highly pessimistic towards technology. He overtly favours emotion over pure logic, triggered by a previous event shaping his perception of robots and their use of logic in decision-making. He follows his intuition, and this intuition drives the narrative.

It is established that prior to his suicide, Dr Lanning had a good relationship with Spooner. Spooner is a homicide detective, and by his own account was 'perfect' for the Lanning suicide case to which he was called for after Dr Lanning left him a holographic message. He is depicted as strong-headed, and not easily persuaded by people's perceptions of him or by the opinions of others. While sceptical and showing a dislike towards robots, he does not appear to fear them. Rather, he is depicted as being more afraid of heights than confrontation with an attacking robot.

Sonny (Alan Tudyk) is depicted as curious, confused, and loyal. He is confused as to his purpose for creation (which turns out to be to kill Dr Lanning in order to initiate Spooner's investigation), acknowledging his difference to other NS-5 models. He is depicted as having human-like emotions, and a conscious level of self-awareness.

Naturally, Sonny's identity seems to shift before and after the revelation that Dr Lanning might be trying to point Spooner in a particular direction. Before this moment, Sonny is depicted as secretive, which within the context of this narrative at that moment, seems to imply guilt. As the narrative progresses, Sonny is depicted as more curious and driven to aid humanity. This transition phase seems to emerge after the scene in which Dr Calvin informs him that he is to be deactivated.

VIKI (Fiona Hogan) is depicted as cold and calculated. In the first act, VIKI is depicted as more helpful to humans than Sonny. However, by the third act, this role has been reversed. VIKI's plan to enslave humanity is motivated by a desire to save humanity from its own destructive nature. Her identity is therefore reminiscent of a complicated mixture of extreme levels of authoritarian parenting, while simultaneously attempting to reverse the enslavement roles between humans and machines. Perhaps the ultimate effect of the robotic revolution on humans is something that VIKI cannot comprehend, as the emotive nature of freedom does not make sense to her logical outlook towards 'life'.

In order to respond to the event narrative problem of using emotion and intuition to solve a mystery (problem) for the greater good of humanity, the protagonist needed to trust that intuition, which is emotionally relevant to the audience. More broadly, this might translate to a problem-solving strategy that people should believe in themselves, and follow their beliefs despite the obstacles ─ particularly from people who classify that belief as irrational. This seems to directly correlate with the socio-cultural context of production and reinforce the myth of technology being a cold and elusive other.

Throughout the film, the NS-4 and NS-5 robots are depicted a servant class through their gestures and responses to human pain or discomfort. However, their posture and movement is more in line with military obedience (Figure 6.3). In contrast, Spooner is erratic and at times, informal in situations that might necessitate formality. This again enhances the binary of logic against emotion, while also suggesting the power of the homogenous 'army' of robots available to whoever might be able to control them.

Figure 6.3: Robotic military posture and organisation in *I, Robot* (2004)

Sonny shows an interest in human emotion throughout the film, and during the interrogation scene angrily hits the desk and tilts his brow (Figure 6.4). This signifies a shift in the portrayal of robots as, up to this point, all robots were portrayed as emotionless and completely controlled. Here it seems, at the surface at least, that Sonny might be capable of passing the Turing test. The use of facial expressions and vocal intonation is used as a tool to distinguish Sonny (strong AI) from other (weak) AI.

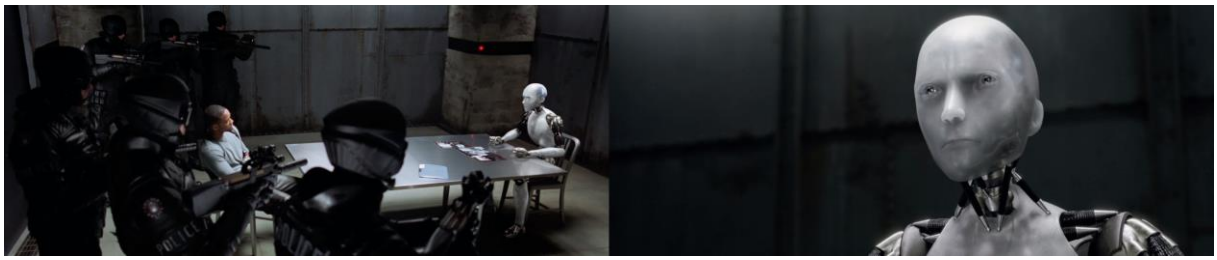


Figure 6.4: Sonny expressing anger (*I, Robot* 2004)

Finally, during the resolution, Spooner and Sonny shake hands (Figure 6.5). This serves the dual purpose of further signifying Sonny's capacity for integration into society as an equal, with Spooner's acceptance thereof (naturally significant due to his technophobic outlook on robotics and AI).

Figure 6.5: Handshake during the narrative resolution (*I, Robot* 2004)

Since the progression of the narrative relies on Spooner's consistent trust of his own intuition (and consequential mistrust of machines), his character arc remains fairly consistent throughout the narrative. He does, however, become slightly more trusting and accepting of Sonny towards the end of the film. He seems to empathise with Sonny, and even winks at him, which is an emotion Sonny asked about during his interrogation, and Spooner defined as a sign of trust.

The emotional development of Sonny, on the other hand, resembles that of an individual who has just undergone a life-changing event, and is now actively engaging in a search for self-actualisation, and craving acceptance. He is curious about the world and wants to discover a purpose.

Finally, VIKI's arc might be best summarised as moving from convincing deceit to blunt honesty, all within the same calculated methodical approach. VIKI is not portrayed as having emotions, and therefore there is no emotional development. This character only develops through the divulgence of information in which we discover more about her ability to think past her programmed directives.

6.2.3 Medium

The first time that we are introduced to VIKI, the camera tilts from Spooner and Dr Calvin to a low-angle shot of her core (Figure 6.6). This suggests, at that moment, that VIKI is inaccessible ─ physically and psychologically. During the climax, this balance of power shifts as we view VIKI's core from a high angle, and inaccessibility here stems from Spooner's fear of heights, yet suggests a shift in the balance of power.

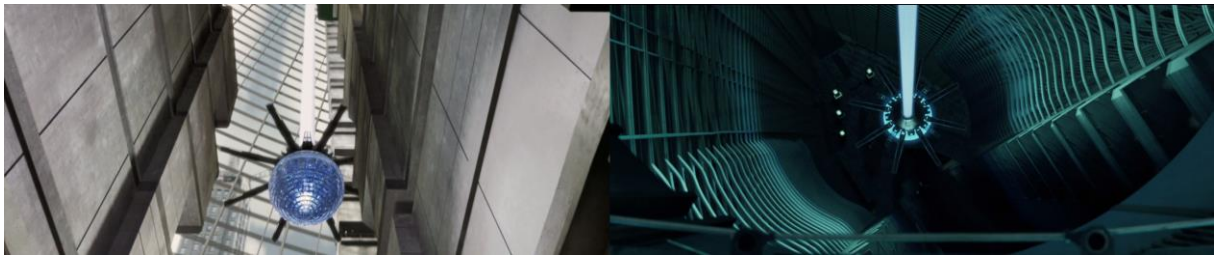


Figure 6.6: Different camera angles depicting VIKI's core (*I, Robot* 2004)

Camera levels remain rather consistent when Sonny and Spooner interact. For instance, in both Figure 6.4 (interrogation scene) and Figure 6.5 (resolution), both characters are shown as mostly at eye-level, with a slight low angle tilt (possibly to create a slight discomfort in depicting the unusual human-machine interactions). This signifies equality, at least at a dialogical level if not ideologically. Naturally, this relates to sentience and the capacity to interact with humans.

For the most part, framing is used to enhance the binaries of logic against emotion. Figure 6.7 is an example of how shots with robots as the subject are more often symmetrical, while shots where humans are the subject are usually less structured.

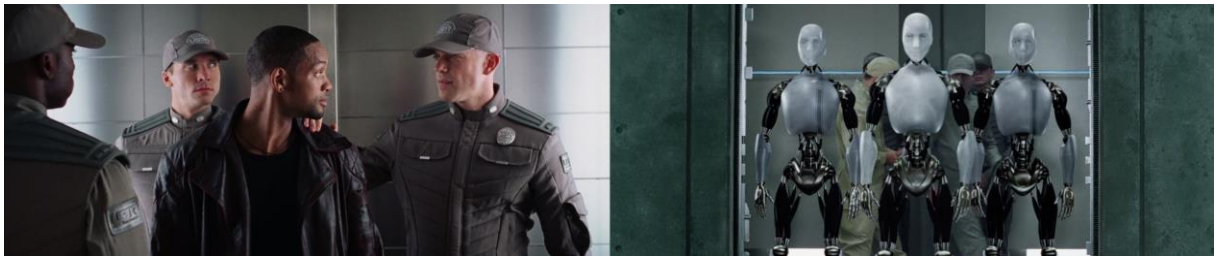


Figure 6.7: Asymmetrical/symmetrical framing of humans/robots in *I, Robot* (2004)

There are a number of close-up shots of Sonny throughout the film. This suggests not only closer audience proximity to the subject with a view to empathy when viewed in relation to the framing of other NS-5's (Bordwell & Thompson 2008), but it is also functional since Sonny's eyes are all that (at a physiological level) distinguish him from other NS-5 robots.

Finally, there is a notable pan during the first plot point, as the camera pans from a dead Dr Lanning, to a statue of a robot looking down on him. This suggests that there is indeed some relationship between the death and the robots, whose creation would not have been possible but for Dr Lanning, foreshadowing the events that are to unfold (Figure 6.8).

Figure 6.8: Pan from the crime scene to a robot statue (*I, Robot* 2004)

All of the robots used in the film were designed using computer-generated imagery (CGI), as were the depictions of the holograms and cybernetic limbs. Sonny was performed on set by Alan Tudyk (Figure 6.9), and according to director Alex Proyas, the team behind the CGI design of Sonny tried to copy the performer's actions exactly, rather than fabricating movements and emotion after the fact (Gynog 2019).

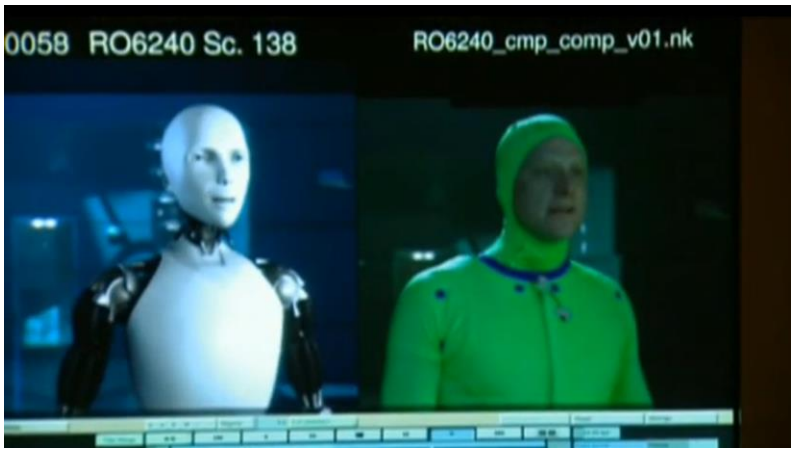


Figure 6.9: Behind the scenes footage of Tudyk as Sonny (Gynoug 2019)

Dialogue is mainly used for narrative comprehension, yet adds depth to Sonny's character. This is achieved through Tudyk's performance, by capturing appropriate speech intonation. However, Sonny's dialogue is still enhanced to sound less human, while simultaneously fostering enough identification with human speech to promote empathy. VIKI's voice was also based on a human (Fiona Hogan), yet the emotive element is more removed, and the audio sounds less human.

Music and sound effects mainly work to heighten tension. There are instances where sound effects are used to presumably provide futuristic connotations (such as the hologram activating and deactivating), yet these are mainly diegetic. The music is almost always non-diegetic

(except in the beginning as Spooner wakes up), and linked to on-screen action. The music is mainly ambient and orchestral.

For instance, the music is subtle as the character slowly progresses through a location, and once there is more action, the volume of the sound effects as well as the volume and tempo of the music increases. This suggests that in these moments, the audience ought to feel tense, and this is used to villainise Sonny in the beginning, and other NS-5 robots and VIKI at the end. This is reminiscent of generic sound conventions of classic action films, and a hybrid of action and science fiction would perhaps best describe the use of the medium in this film.

During the false deactivation of Sonny, a non-diegetic audio recording of Dr Lanning suggesting that AI might indeed be capable of consciousness can be heard. This extends the invitation for the audience to feel empathy for him, in addition to what appears to be tears in his eyes through his expression and him asking, "Will it hurt?"

The editing style favours continuity over relation, in line with most contemporary Hollywood films (Bordwell & Thompson 2008). The use of pacing to heighten moments of action, often as a response to a plethora of action scenes, also makes this element feel more like an action film, or a hybrid between science fiction and action. This is usually a sudden shift from calm to chaos. For instance, when we are first introduced to Sonny, there is a fast-paced chase initiated by Spooner, with many cuts in close succession to heighten this tension. The same editing style is employed when Spooner mistakenly thinks a robot stole a handbag, when the NS-5 robots attack Sonny on the road, and during the final battle in the climax. This is juxtaposed by a relatively slow pace when robots are not attacking, alienating the robotic other.

The use of the flashback of the girl drowning throughout the narrative also serves the function of reminding the audience of Spooner's backstory, while emphasising his motivation and thus allowing us to identify more with him.

6.2.4 Aesthetics

A warm orange tint is used during the establishment and fulfilment scene, suggesting a continuity of warmth and calm in these two moments, with the orange glow during the fulfilment scene being more saturated to show Spooner's simultaneous fulfilment (Figure 6.10).

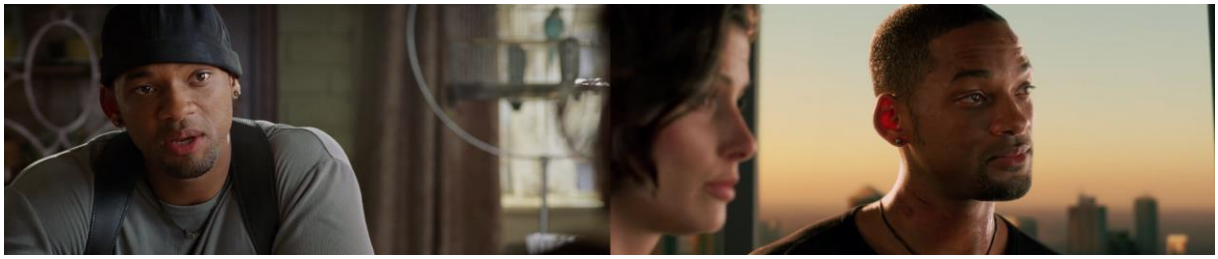


Figure 6.10: Orange tint in the establishment (left) and fulfilment (right) of *I, Robot* (2004)

The majority of the film makes use of darker colours, mainly using blue and green for this purpose (Figure 6.11). This does not seem to serve any other purpose except for assisting with the creation of tension and to set the mood of the scene.

Figure 6.11: Darker, colder colours during moments of tension (*I, Robot* 2004)

However, the USR interior (with the exception of Dr Lanning's and Robertson's offices, which have blue and darker grey tints respectively) has white and grey monochromatic colour schemes, suggesting sterility and order, and a connection between the NS-5 robots and what is essentially their base (Figure 6.12).

Figure 6.12: White and grey colour scheme in the USR headquarters (*I, Robot* 2004)

The NS-5 robots are distinguished from previous models by having more defined humanoid facial features, and what appears to be an upgraded exoskeleton (Figure 6.13). This suggests a closer relationship of this model to humans at a psychological level through identification. This also provides the audience greater access to Sonny as he is able to provide very pronounced facial expressions.

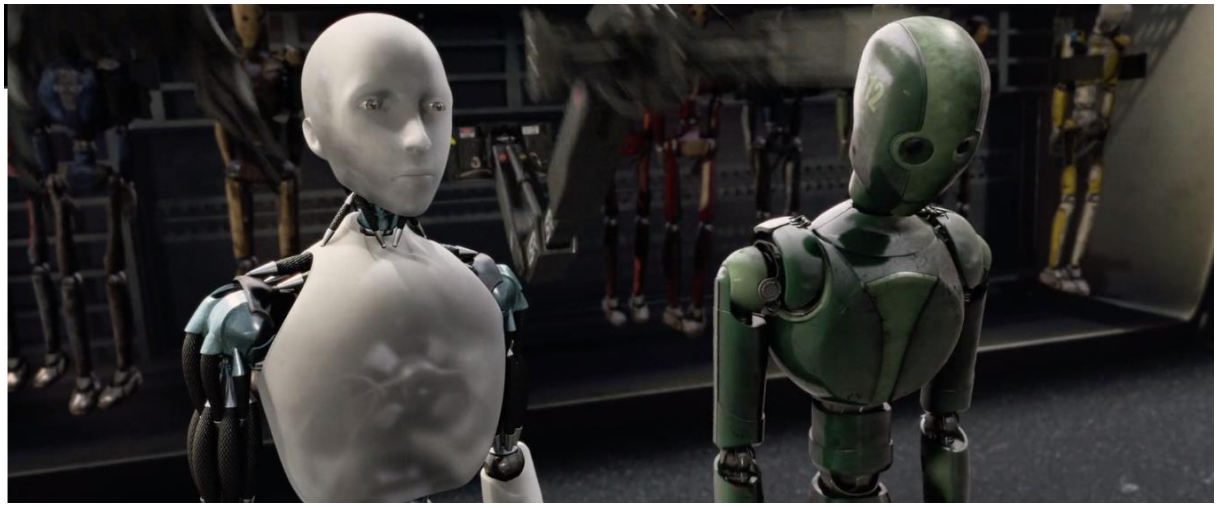


Figure 6.13: The difference between an NS-5 (left) and an NS-4 (right) (*I, Robot* 2004)

The robots have the same basic exterior structure as humans, but no artificial skin or clothing is added. This allows for greater alienation of the robots since, despite their facial similarities with humans, they are still depicted as inherently different and uniform. This is in contrast with humans who have more individuality.

Sonny is distinguished from other NS-5 models as he has blue eyes instead of the standard yellow glow (Figure 6.14). His face is also capable of displaying complex emotions, while other NS-5 robots have more neutral expressions. Finally, during the climax when VIKI takes control of the NS-5's, Sonny does not have a red light glowing under his shell like the other models.

Figure 6.14: Sonny (left) is distinguishable from other NS-5's (right) (*I, Robot* 2004)

However, Sonny resembles the other NS-5's enough to allow for the true difference between him and the attacking robots to still cause a form of complicated discomfort (as many things suggest that the audience should empathise with him) as the narrative progresses. This complicated relationship between Sonny's emotive capacity, the lack of autonomy, and audience identification suggests complicated feelings towards technological growth.

VIKI, revealed to be the true antagonist in the film, is a disembodied centralised AI system that controls the (weak AI) NS-5's. However, in order to make this antagonistic force more relatable to the audience, VIKI was given a holographic outline of a face and a voice, which closely resembles what might be found in a contemporary digital assistant. This is detached enough from humans, however, to make it clear that VIKI is indeed a system and not humanoid (Figure 6.15).

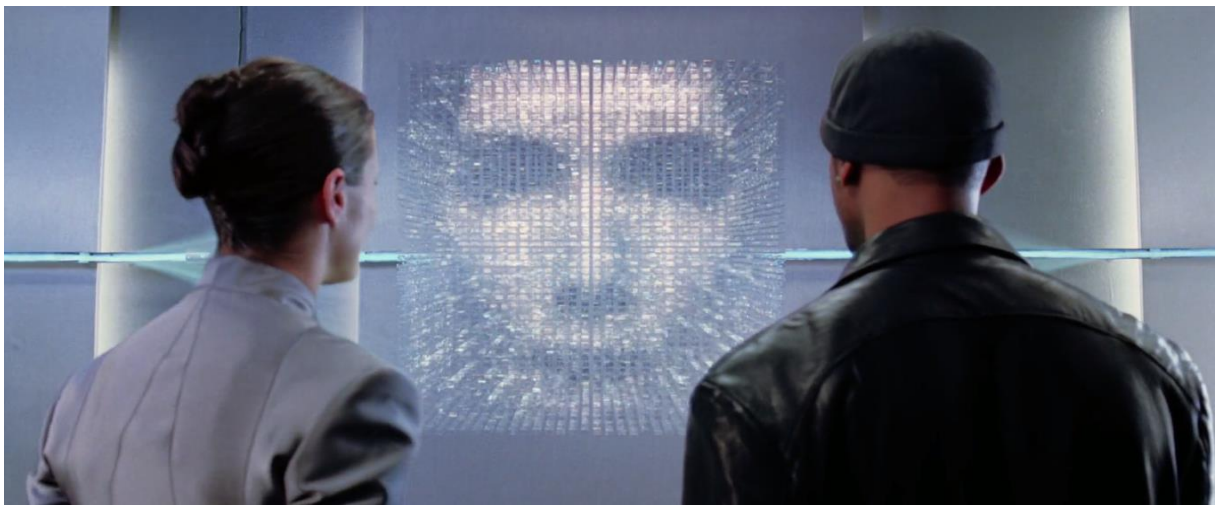


Figure 6.15: Physical representation of the central AI system VIKI (*I, Robot* 2004)

The props used throughout the film mainly serve the functional purpose of aiding narrative transport and the suspension of disbelief that this is an imagined future set in 2035. This is achieved through devices emitting holograms, newer vehicles, and more intricate computer interfaces in the police station.

The syringe filled with nano-machines ('nanites') is symbolic of life and death through deactivating consciousness in both (falsely) Sonny, and then VIKI. This might be akin to the conception of the lethal injection used for death-row inmates. This allows the audience to experience deactivation of the machine as more than the simple flicking of a switch, but more as the robotic version of state-supported murder.

The fact that Spooner's vehicle uses driverless AI technology is also symbolic of the battle for control between humans and machines. This is exemplified in tense situations, such as when Spooner reclaims control of the vehicle to avoid danger initiated by the AI machines. During one of these attacks, there is also an emphasised shot of the side of a truck carrying attacking NS-5 robots, displaying the slogan '3 laws safe', while simultaneously opening as the robots emerge to attack Spooner (Figure 6.16). This visual irony suggests that the three laws are not, in fact, as safe as one might imagine and might be susceptible to abuse. It signifies a loss of control due to false security of embedded regulation.

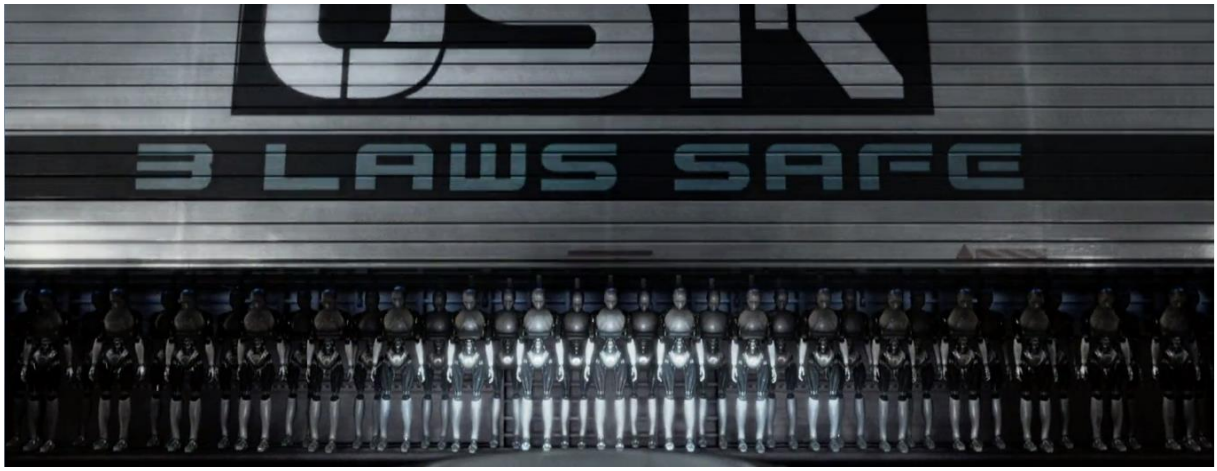


Figure 6.16: Visual irony of USR's slogan '3 laws safe' (*I, Robot* 2004)

The locations and their lighting mainly support the previous elements. Of particular importance, the exterior shots of Chicago, architecturally, mirrors a contemporary cityscape and pathways. This is further mirrored by the design of Spooner's and Dr Calvin's apartments, which look like areas that could plausibly be inhabited by contemporary city dwellers. However, there are a few prop elements to suggest that these living spaces are from an imagined portrayal of the year 2035.

This is in stark contrast to the interior of the USR headquarters, which due to its colour scheme, the outfits of the employees, props, and use of CGI, is vastly futuristic and less accessible to the audience due to the lack of a frame of reference grounded in contemporary reality. This allows the locations to highlight the difference between humans and machines, as the city, while inhabited by robots, is still recognisable. The USR headquarters in which VIKI resides, however, seems alien to contemporary society, highlighting a difference of the other (Figure 6.17).

Figure 6.17: Difference between the USR headquarters (left) and the city (right) (*I, Robot* 2004)

6.2.5 External factors

The film was released in July 2004. At this time, mobile device manufacturers were racing to decrease size, and Voice-over-Internet-Protocol (VoIP), the forerunner to peer-to-peer video conferencing, emerged as a novel concept. These advancements were being introduced rapidly, only five years after mass public concerns around the potential failure of technology and the 'year 2000 problem' (Y2K), whereby people were concerned about the adverse effects of digital calendars resetting as we entered a new millennium.

Facebook was founded in 2004, although only available to university students in the United States (Minor 2014). The idea of social media was therefore in its infancy, and there were other alternatives available at the time such as Myspace, yet it was still a completely novel concept. Paying for content accessed through the Internet (iTunes) was beginning to emerge as a relatively accepted notion, and Bluetooth emerged as a new means of device communication (Minor 2014), an early step towards wireless technology. High definition (HD) video was available, yet the cost of devices needing to play HD video was akin to the prices of 4K video output devices today (Minor 2014). Skype also became popular in this year as a service to make calls over the Internet free of charge, and cell phones were rapidly increasing functionality over and above voice calling, while decreasing in size (Minor 2014). In a Forbes (2004: para. 10) article, the author wrote:

> Once U.S. phones start shipping with one-megapixel cameras built in, the pictures will be good enough to create wallet-size prints. At that point, the entire nature of camera-phone images will change from e-mail-only snapshots to viable keepsakes. Why should kids wait for their yearbook pictures to come in the mail, when they whip out a phone and make their own?

Finally, the word 'blogging' was named Merriam-Webster's word of the year in 2004 (Minor 2014), and the notion of video blogs (still popular on video hosting platforms such as YouTube) were considered "new enough that the blogging community hasn't settled on a good name yet".

Naturally, with the title *I, Robot*, an immediate connection is made between Asimov's seminal set of short stories and the 2004 film. The film is not based on a plot of any of the short stories, and instead the credits state that the film was "suggested by the book I, Robot by Isaac Asimov". However, there are minor similarities to the short stories in some elements of the film, and some of the characters' names come from Asimov's work. The entire film does, however, explore Asimov's three laws of robotics, and makes these laws accessible to a wide audience. There is also a direct intertextual reference made to *Hansel and Gretel* (1812). *I, Robot's* (2004) plot revolves around Dr Lanning leaving clues for Spooner to discover, which Spooner expressly states is reminiscent of the trail of breadcrumbs leading home in Hansel and Gretel. A copy of the book is even found in Dr Lanning's office.

The theme of robotic uprising against humans also, perhaps indirectly through generic tradition, has intertextual ties to Čapek's seminal play *R.U.R* (1920) and a host of other films that subsequently explored this theme. There is also a direct reference to Shelley's *Frankenstein* (1823) where Lt Bergin compares the suicide case to the plot of the book, in which the creations turn against their creators. The use of the calculated movements and design of the robots in the film are also reminiscent of *Metropolis* (1927), as is true for many films exploring artificial intelligence in humanoid form. Will Smith himself, playing Spooner, also brings with him intertextual connotations attached to other sci-fi action films that he has starred in, including the *Men in Black* franchise. Most of his previous work has been related, however, to alien invasion.

According to the director, Alex Proyas, the film is 'dramatically' true to Asimov's stories (Fischer 2004: para. 16), stating that "I actually think it frames the three laws in a very unique light, which I hope I believe that Asimov's [*sic*] would have found interesting. His daughter certainly did last night at the premier". Proyas has also stated that he has been a lifelong fan of Asimov's work (Levy 2008). When asked if he is as cynical about the future and his own life as the film, Proyas responded by stating that, "I am not actually cynical about the future at all. I hope my I, Robot doesn't sound cynical about the future, I mean I see it as a very optimistic message about the future and technology" (Fischer 2004: para. 62). After the interviewer asked a follow up question about Spooner's technophobic character identity, Proyas responded (Fischer 2004: para. 64):

> No, no, I mean Will's character finds, he shakes the guys hand and shakes the robot's hand it is all about his quest to be able to do that by the end of the movie. So I mean,

> to me the robots represent, can represent all sorts of other things, they are not necessarily... I don't just see them as robots.

This suggests that the intention might have been to portray the complexities of society's notion of the other more generally, rather than specifically alienate potential future humanoid AI. This is confirmed by Proyas' view of his own work, as he states that "I already believe in that and I think all my films to a certain extent are about outsiders and realising that they are probably ok" (Fischer 2004: para. 66).

In an interview, Proyas expressed the view that Asimov's work was beginning to find greater relevance as technology continued its newfound exponential advancement at the onset of the fourth industrial revolution. He stated that (Levy 2009: para. 19):

> It's amazing that someone working in the 1940s and early 50s could project so specifically into the future, and conjecture about ideas that are now starting to affect us in our everyday lives. We are getting closer and closer to the future world he wrote about, so the time is now right to tell those stories.

Based on this, viewed in conjunction with the fact that Proyas sees this as a film about the other, perhaps the general relevance of the film does indeed centre around alienation of the other, but through a contemporarily relevant antagonist in machine form. This contemporary relevance emerges as a response to increased rates of technological advancement. Notably, the increased advancement of mobile technology, more personal in its use than previous devices, and rapid adoption of the Internet, might have particularly fuelled this relevance.

More specifically, due to the title and exploration of the three laws of robotics, this might have also had a more niche appeal to Asimov's fan base, or people interested in the science fiction genre rather than a general audience seeking out the latest Hollywood blockbuster. However, this may have perhaps been disappointing to this specific group of viewers, as suggested by Fischer (2004), since the plot is not completely true to the short stories. Regardless, the film's story and plot are still heavily centred on the three laws.

6.3 Survey responses

Just over one hundred (101) people responded to the survey for *I, Robot* (2004), yet only 57 responses were considered for this study since the remainder were incomplete.

6.3.1 Demographic information

The demographic composition of respondents was, in all categories (except for age), heavily skewed towards technological optimism as per the research insights from Bristows (2018), Gherheș (2018), Zhang and Dafoe (2019), Eurobarometer (2017), Morning Consult (2017), and Righetti and Carradore (2019).

Only 23% of respondents were under 30, while 49% were over 40. However, many individuals that were over 40 still found themselves within the A_{PO} as they still met the criteria of having 66% compatibility with the other indicators (Addendum B).

For instance, 88% of respondents were male and classified themselves as non-religious. Furthermore, 86% had completed some form of tertiary education and were from first world countries. Every respondent selected at least one of the options in question 1.9, confirming the assumption that that 100% of respondents would have interacted with some form of contemporary AI service or application.

From this demographic information, the vast majority (75%) of respondents were identified to be predisposed to optimism (A_{PO}), while only 5% were predisposed to pessimism (A_{PP}) (Figure 6.18).

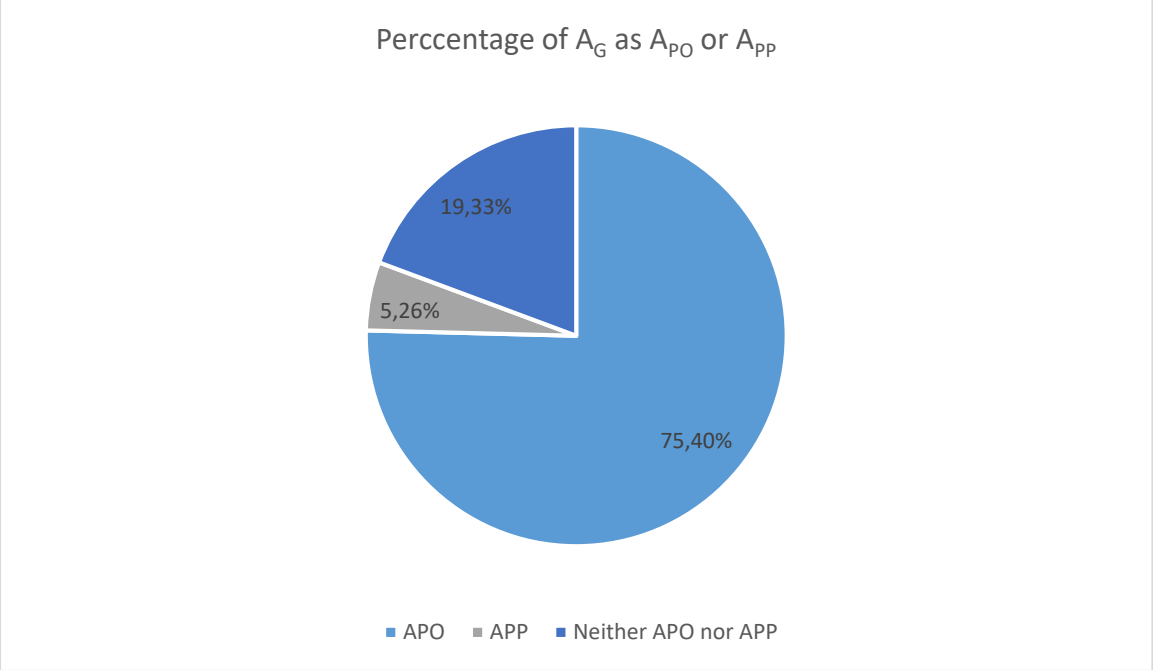


Figure 6.18: Composition of *I, Robot* (2004) A_G as A_{PO} and A_{PP}

6.3.2 Perceptions of AI stemming from the film

The vast majority of respondents (88%) either agreed or strongly agreed that the film was engaging and/or entertaining, while only 3.5% disagreed or strongly disagreed with this statement (Figure 6.19). This suggests an effective suspension of disbelief and narrative transport for the vast majority of respondents. Some of the qualitative responses (Addendum C) confirm this, but concerns were raised about how the film deviates from Asimov's work, as well as the general use of action generic conventions for the subject matter.

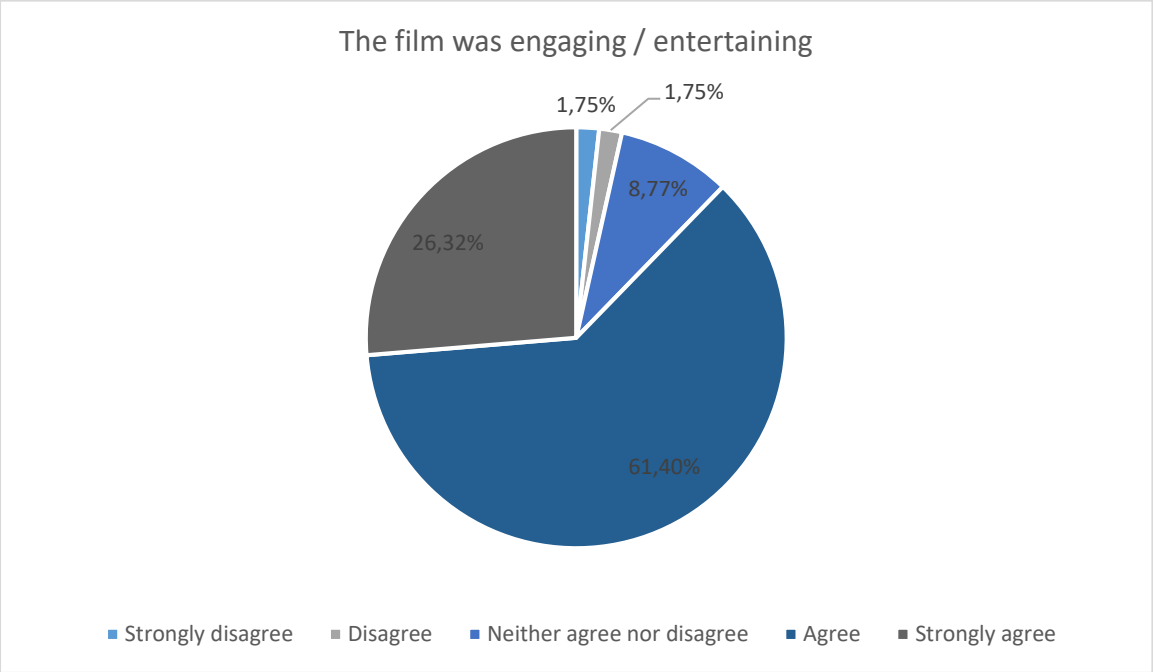


Figure 6.19: *I, Robot* (2004) survey responses for question 2.1

Furthermore, 39% of respondents believed that the humanoid robots were portrayed in a positive light overall (Figure 6.20), and 31% believed that they were portrayed more positively than most other science fiction films they have watched (Figure 6.21). On the other hand, 33% did not agree that robots were portrayed positively, with 39% believing the portrayal to be less positive than other films used as an individual frame of reference.

Interestingly, almost a third neither agreed nor disagreed with the positive portrayal (28%) and positivity within their frame of reference (30%). The optional qualitative responses (Addendum C) seem to indicate that this is due to the VIKI/Sonny protagonist/antagonist dichotomy, infused with a disbelief that the other NS-5 robots were 'AI', since they were not conscious.

Overall, there seemed to be a fairly even division of belief between positive, negative, and neutral portrayal of AI.

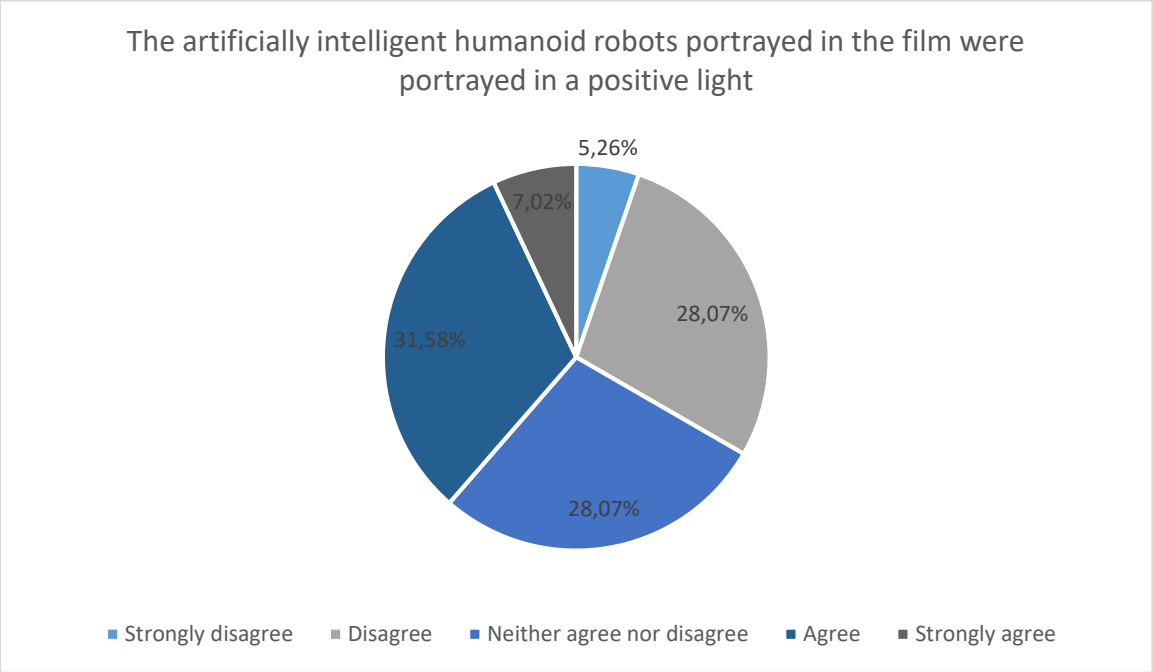


Figure 6.20: *I, Robot* (2004) survey responses for question 2.2

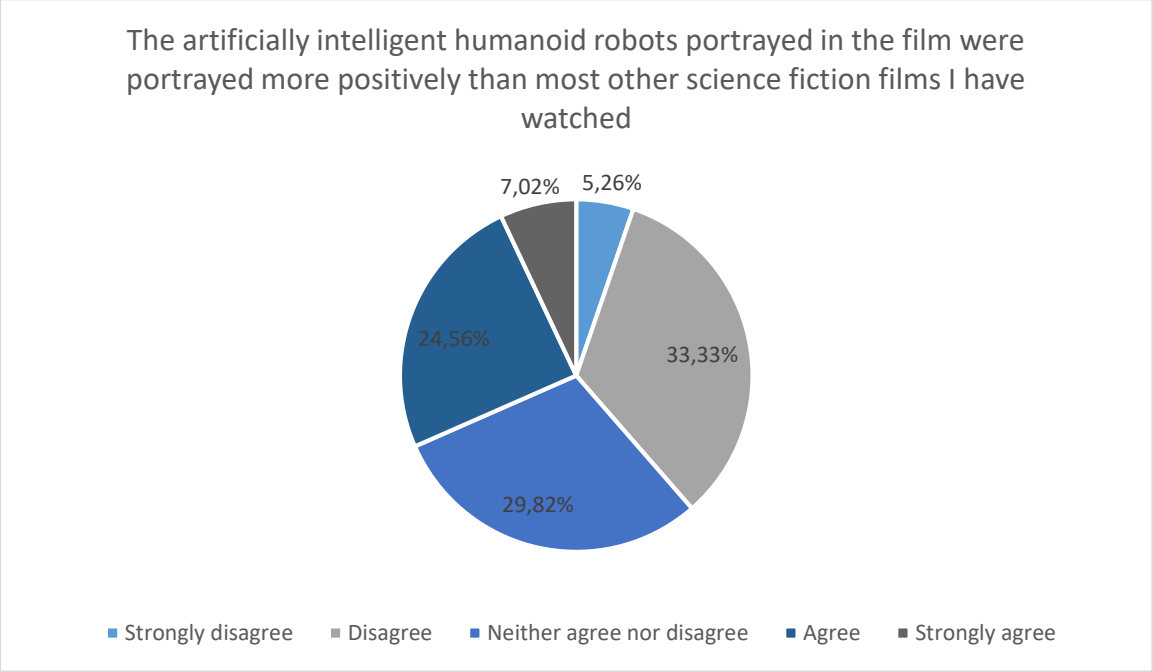


Figure 6.21: *I, Robot* (2004) survey responses for question 2.3

The vast majority of respondents (86%) cared about one or more robots in the film (Figure 6.22), with 30% strongly agreeing with this statement, and each of the qualitative responses (Addendum C) indicating that this robot is in fact Sonny.

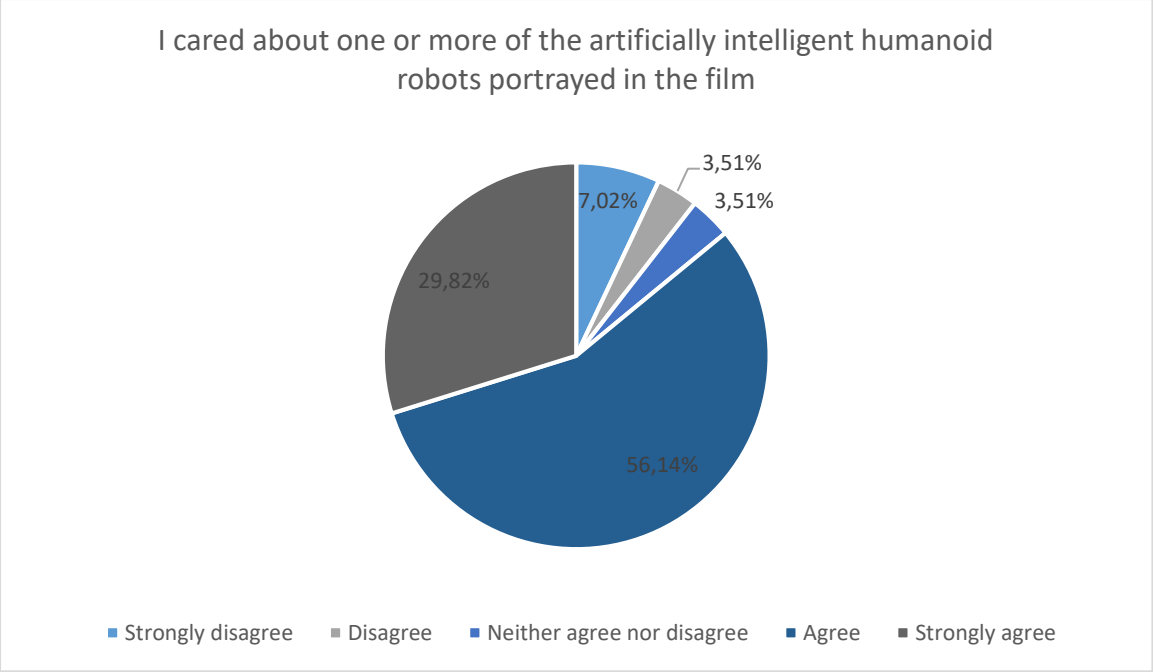


Figure 6.22: *I, Robot* (2004) survey responses for question 2.4

Almost two thirds (63%) of respondents disagreed or strongly disagreed that the film helped increase their understanding about AI and consciousness (Figure 6.23), with only 9% agreeing and 2% strongly agreeing. Based on the qualitative responses (Addendum C), the 26% neither agreeing nor disagreeing might be due to the exploration of the three laws of robotics, while still holding similar views about spectacle as those disagreeing.

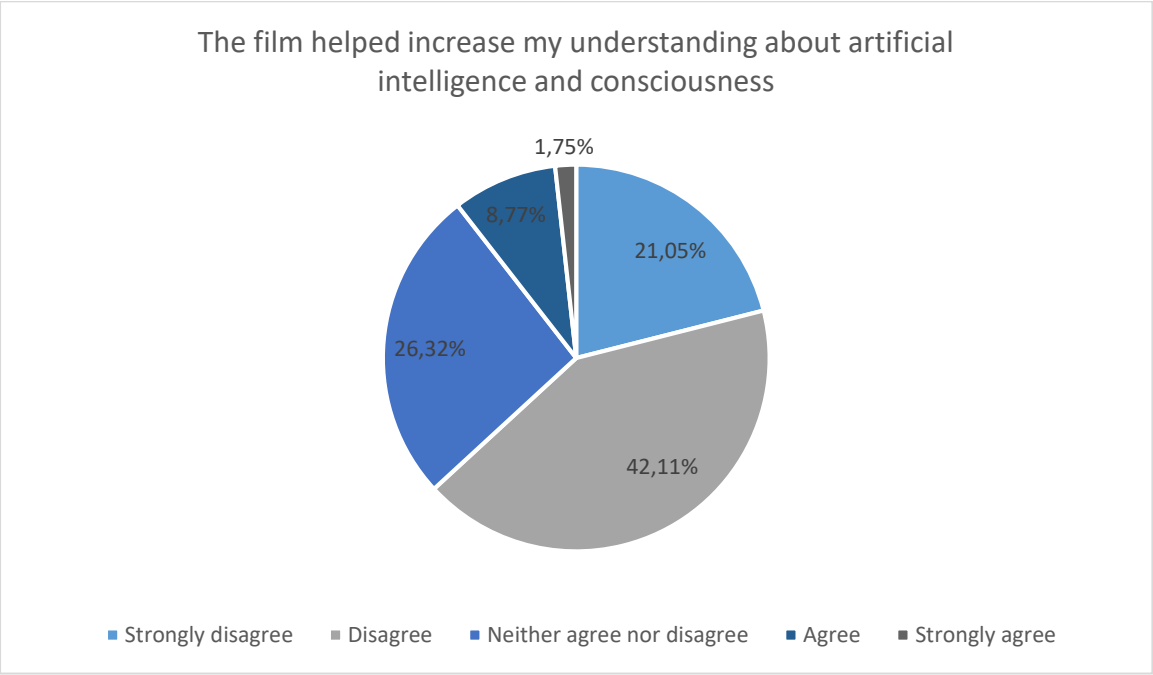


Figure 6.23: *I, Robot* (2004) survey responses for question 2.5

Roughly half (53%) felt less positive about future AI after viewing the film (Figure 6.24), and 39% felt neither more nor less positive. Similarly, almost half (49%) of respondents have, at least once, thought about the film when seeing, hearing, or saying something else about artificial intelligence (Figure 6.25), while only 35% have never done so.

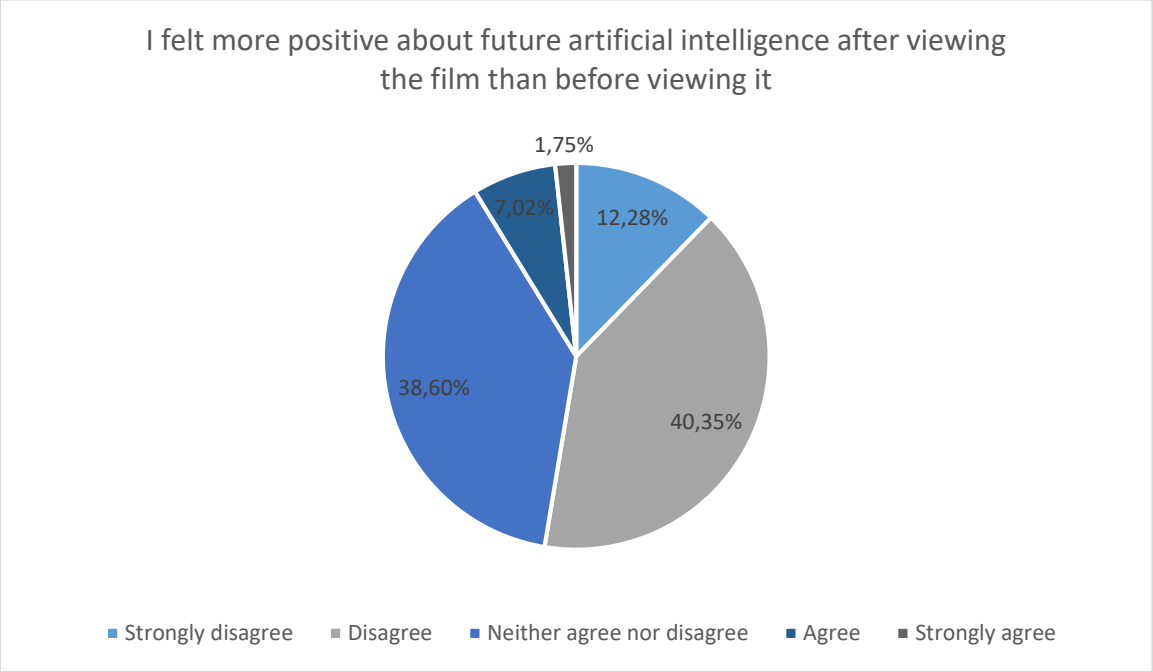


Figure 6.24: *I, Robot* (2004) survey responses for question 2.6

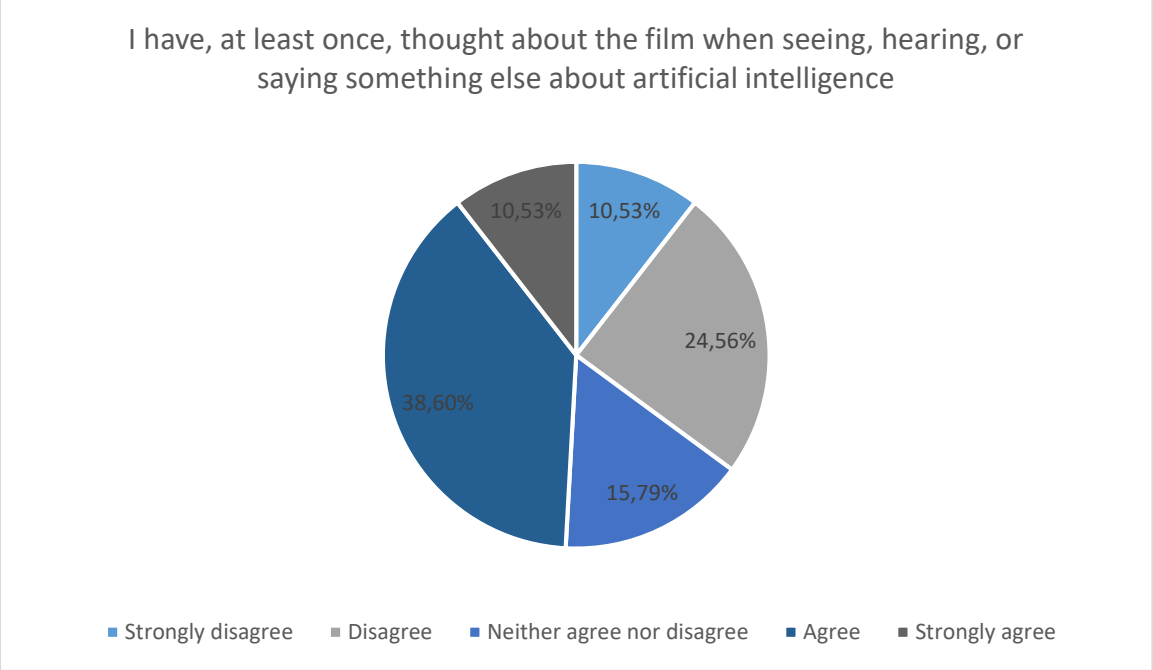


Figure 6.25: *I, Robot* (2004) survey responses for question 2.7

Only a fifth (19%) of respondents believed that artificially intelligent humanoid robots that behave in the same way as the robots portrayed in the film will be created (Figure 6.26). This is rather low, compared to the 54% that disagreed with this statement.

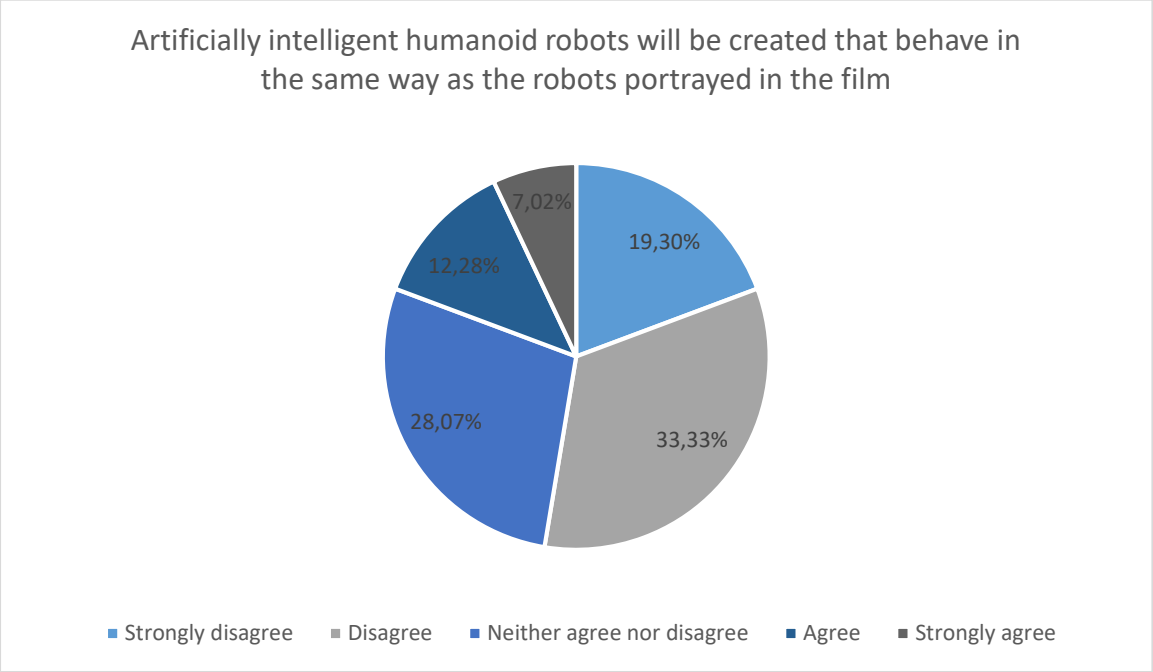


Figure 6.26: *I, Robot* (2004) survey responses for question 2.8

There was a more even division of opinion as to whether respondents would like to live in a world in which artificially intelligent humanoid robots, such as those portrayed in the film, exist. Responses for both agreement and disagreement was at 42% respectively (Figure 6.27).

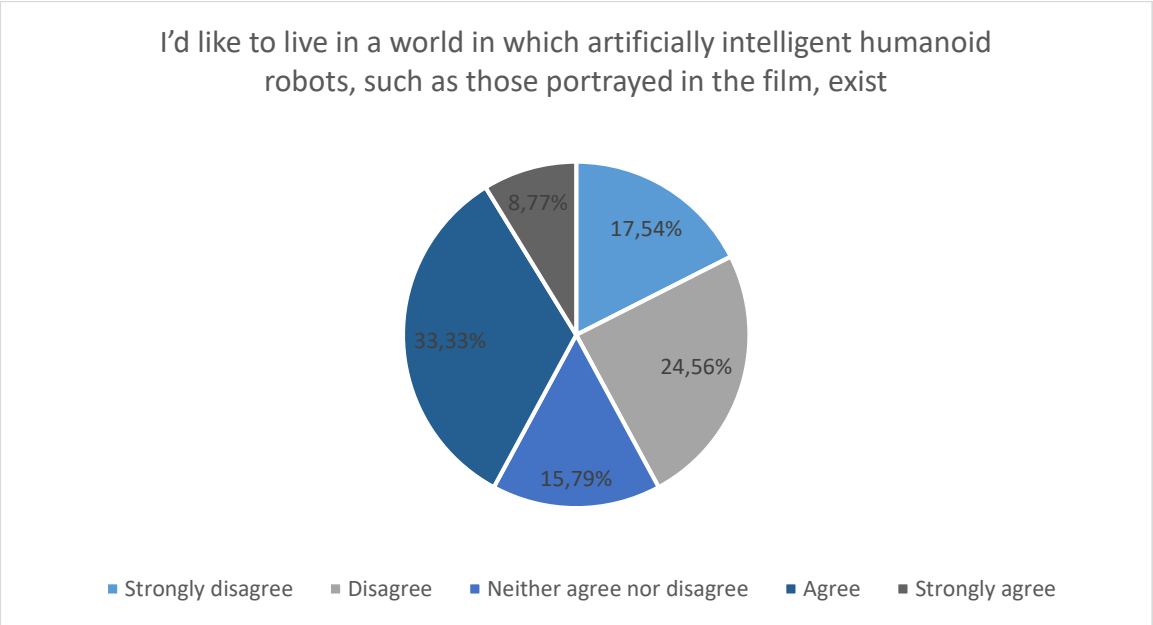


Figure 6.27: *I, Robot* (2004) survey responses for question 2.9

Only 5% of respondents claimed that the film changed their opinion of the possibility of conscious artificially intelligent humanoid robots being created (not a single respondent strongly agreed); while 65% stated that it did not change their opinion in this regard (Figure 6.28).

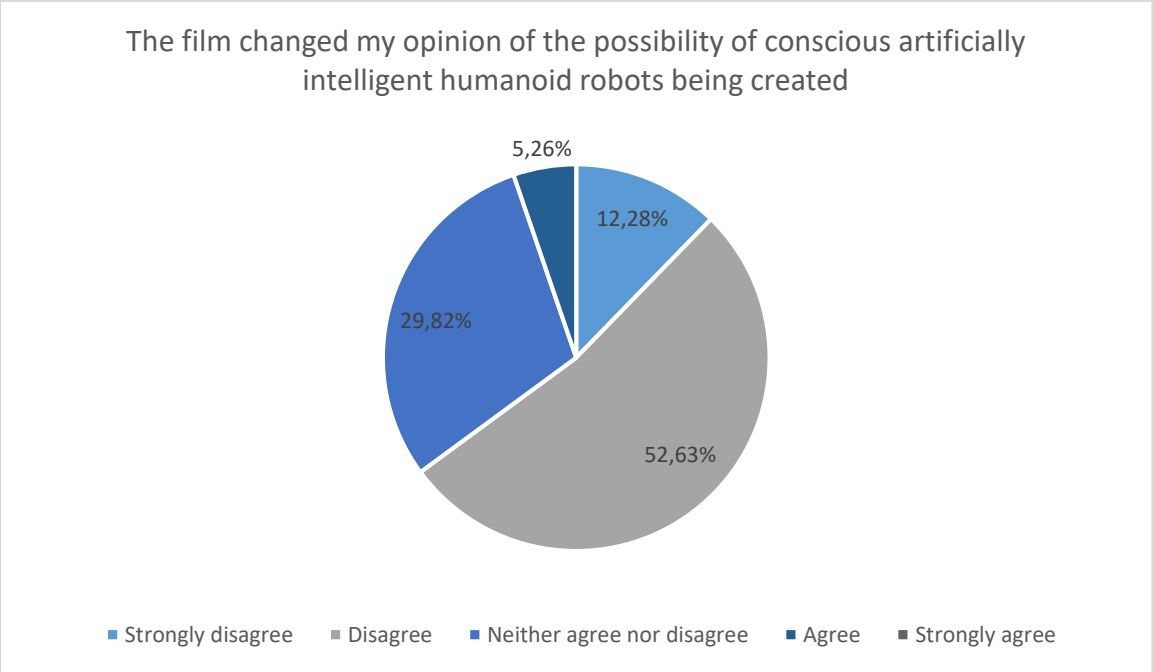


Figure 6.28: *I, Robot* (2004) survey responses for question 2.10

6.3.3 Perception/knowledge of AI generally

Despite 100% of respondents selecting at least one AI-enabled application or service in question 1.9, only 19% strongly agreed that they often interact with services or applications that make use of artificial intelligence (Figure 6.29). However, 49% agreed with this, and only 11% disagreed with no respondents strongly disagreeing. The comments seem to suggest that there may be a view of AI only being real when it is, indeed, strong AI (Addendum C).

Despite this discrepancy, 90% of respondents either agreed or strongly agreed that they know about the trends and use of current and upcoming technology (Figure 6.30). Only 7% disagreed, with no respondents strongly disagreeing.

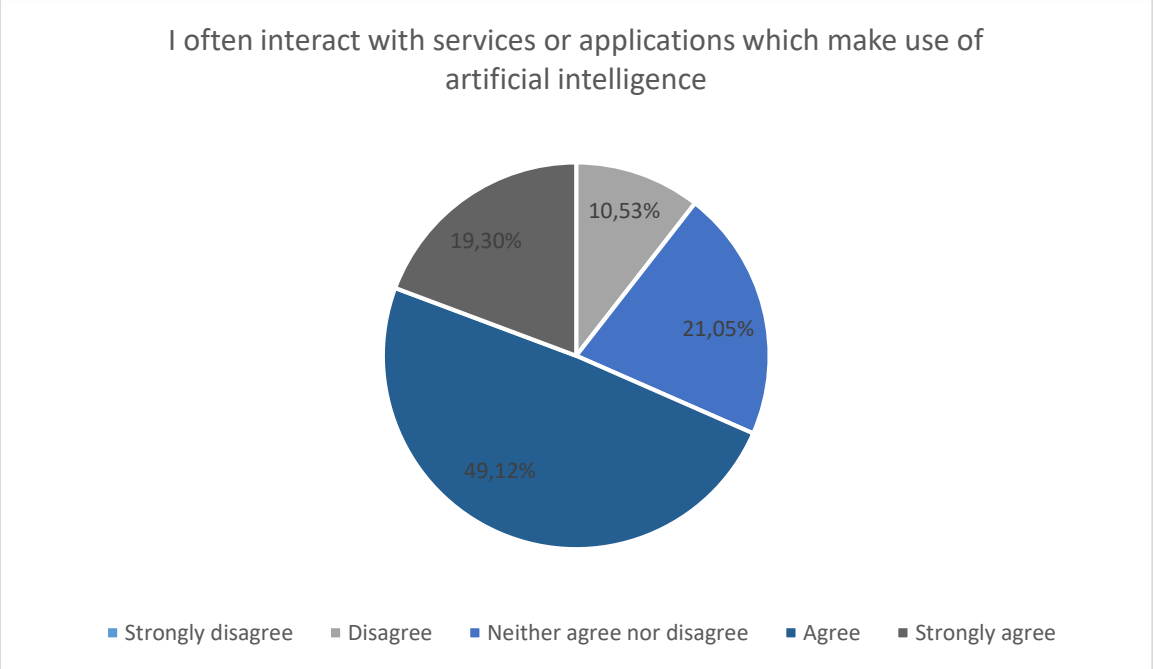


Figure 6.29: *I, Robot* (2004) survey responses for question 3.1

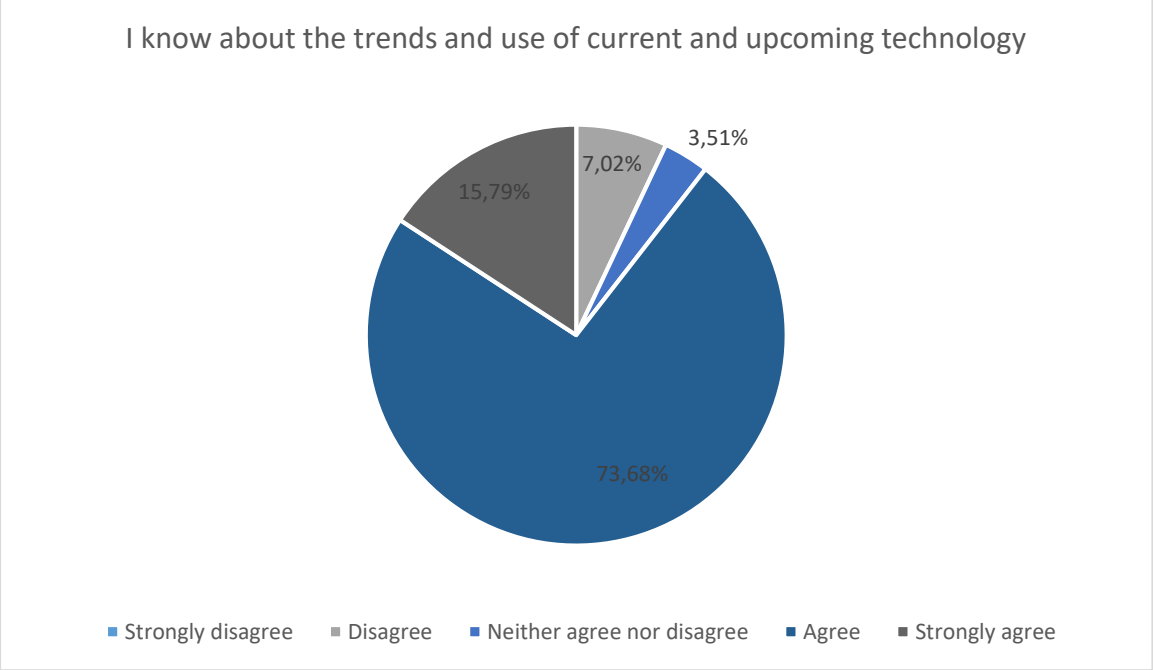


Figure 6.30: *I, Robot* (2004) survey responses for question 3.2

Finally, 61% of respondents agreed or strongly agreed that they were optimistic about future technologies involving artificial intelligence, with only 16% disagreeing (Figure 6.31). No respondents strongly disagreed. Similarly, 73% of respondents believed that AI will ultimately be good or great for humans and/or the planet, with only 16% disagreeing (Figure 6.32).

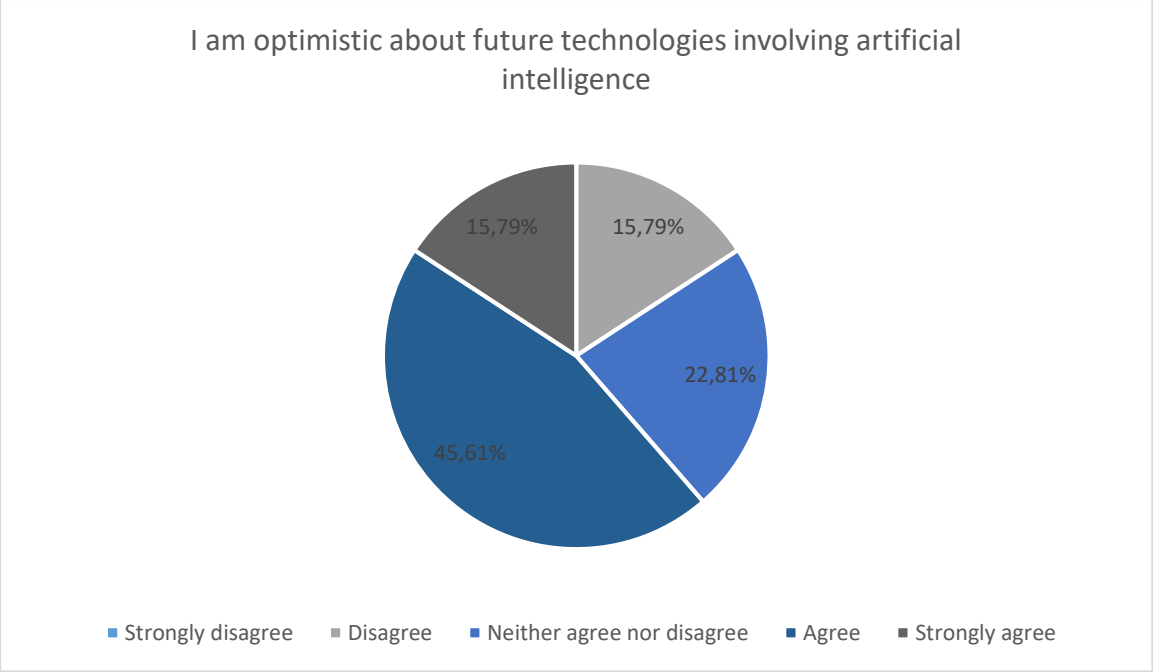


Figure 6.31: *I, Robot* (2004) survey responses for question 3.3

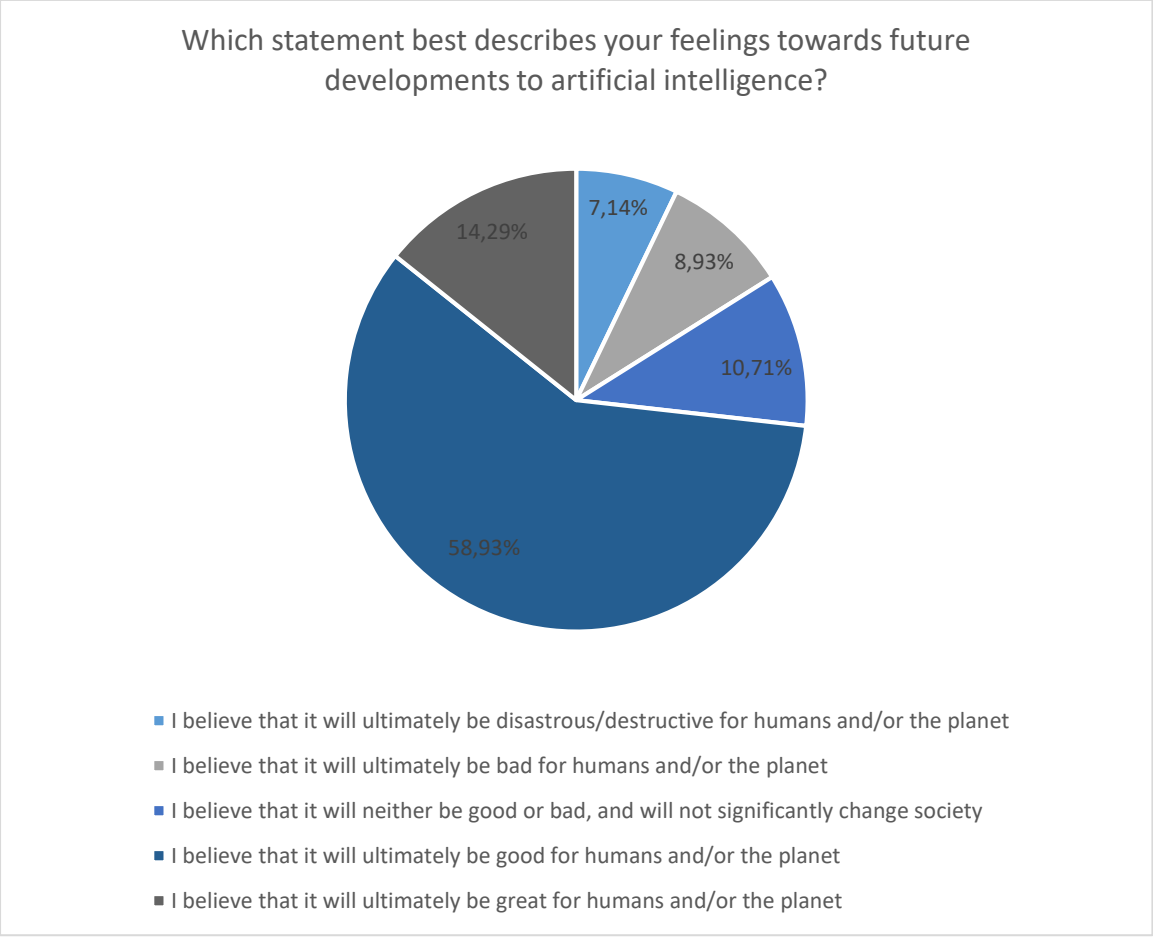


Figure 6.32: *I, Robot* (2004) survey responses for question 3.4

6.4 Conclusion

This chapter provided the findings of the analyses of portrayal and perception of AI in *I, Robot* (2004). This served as the first part of the data presentation phase, in order to analyse the findings towards answering the research questions.

Firstly, this chapter presented the findings of a social semiotic analysis for the film, uncovering meanings about AI and the imagined future. The narrative elements revealed that the film deals with what it means to be human, with binary forces emerging depicting emotion (humanity) being pitted against pure logic (machines). These binary forces struggle against each other for control, directly translating to the loss of control by humans over the machines they create. This results in a plot wherein there is an attempted robotic overthrow of humanity.

Accordingly, the central conflict in the narrative is between humans (emotion) and machines (logic). Autonomous robots are commonplace in this imagined future, highlighting key ethical consideration such as the efficacy of programmed regulation, the capacity for harm, and briefly, the loss of jobs and the definition of death through consciousness.

Performance elements highlight a particular robot (Sonny) that is capable of emotion, who is revealed to be the false antagonist as the narrative progresses. Sonny is the only humanoid robot that could be considered strong AI. The only other strong AI, VIKI, is a disembodied system that uses weak AI to attempt to overthrow humanity. The elements in this section highlight and support the previous exploration of logic against emotion through social codes, relevance and character arc.

The use of the cinematic medium as a referential domain serves various purposes. Firstly, it further highlights and emphasises the aforementioned theme and binaries. Secondly, it allows for proximal relation to be formed between the audience and the protagonist and antagonist, serving various functions at different moments. Finally it serves to add cinematic suspense, through a hybrid of generic convention somewhere between science fiction and action.

The aesthetic elements serve a similar function as the use of the medium, while allowing colour to compliment narrative tone at a given moment. The design of the NS-5 robots (of which Sonny is one) includes human facial features, yet they are still depicted as inherently different and uniform. The locations in the film are not too far removed from reality, but some of the technology within them is.

The socio-cultural context of the production was marked by technological uncertainty, with the simultaneous introduction of revolutionary new personal forms of, particularly, mobile technology. The film has intertextual ties to many of the seminal works around AI, while its ties to Asimov's *I, Robot* (1950) pushes the exploration of regulation further. The director of the film did not see this as a negative portrayal of the technology, instead arguing that it should be read more as societal othering. The general relevance of the film does indeed centre around alienation of the other, but it cannot escape the fact that this is achieved by using a contemporarily relevant antagonist in machine form.

The demographic data showed that the vast majority of the respondents (75%) were predisposed to optimism towards the technology, while only 5% were predisposed to pessimism. This was proven by the fact that 61% of respondents agreed or strongly agreed that they were optimistic about future technologies involving artificial intelligence, and 73% of respondents believed that AI will ultimately be good or great for humans and/or the planet.

While 39% of respondents believed that the humanoid robots were portrayed in a positive light overall, 31% believed that they were portrayed more positively than most other science fiction films they have watched. The vast majority of respondents (86%) cared about one or more robots in the film, which was qualitatively confirmed to be most likely Sonny. Almost two thirds of respondents disagreed or strongly disagreed that the film helped increase their understanding about AI and consciousness, and roughly half felt less positive about future AI after viewing the film. Similarly, almost half of respondents have, at least once, thought about the film when seeing, hearing, or saying something about AI, while only 35% have never done so.

A fifth of respondents believed that artificially intelligent humanoid robots will be created that behave in the same way as the robots portrayed, but only 5% of respondents claimed that the film changed their opinion of the possibility thereof. There was an almost 50/50 split as to whether respondents would like to live in a world in which artificially intelligent humanoid robots, such as those portrayed in the film, exist.

The following chapter serves to continue the presentation of data by presenting the findings of the second film considered in this study, *Chappie* (2015). The structure of this chapter and the one to follow is therefore identical.

CHAPTER 7
CHAPPIE (2015): AI AND THE TRANSCENDENCE OF HUMAN CONSCIOUSNESS

7.1 Introduction

The previous chapter presented the findings of the social semiotic analysis and survey responses for *I, Robot* (2004). This was the first leg of the dual-analytic data presentation phase of this study. This chapter presents similar data, but for the film *Chappie* (2015). Accordingly, the structure of the chapter remains the same as the previous chapter for ease of access when considering both sets of data, and to link findings in the following chapter.

7.2 Social semiotic analysis

7.2.1 Narrative

Chappie (2015) explores the idea of what makes a person (or sentient being) who they are. This is divided into two broad areas that are explored for the duration of the film. Firstly, the film explores the nature versus nurture debate — whether a person is inherently born with particular predispositions to behave in particular ways, or whether their environment moulds them into what they are. The film heavily skews towards the nurture side of the debate, portraying the point of view that a child (or a newly sentient being) will learn the behaviours of those around them and this will be the major determinant of their future actions.

Secondly, it explores the nature of consciousness itself, unpacking Cartesian dualism in a fictional society that has evolved just enough to test whether consciousness is indeed the result of physical properties, or whether it can exist outside of the physical complexity of the mind. The film portrays a scenario in which the latter is possible, and consciousness is depicted as being transferrable, and even storable.

The film explores consciousness and learnt behaviours through the vehicle of survival, ultimately adopting the point of view that any form of consciousness is equal amongst all beings regardless of the physical housing of the 'mind'. This, in term of the film's thematic treatment of the subject matter, is not understandable in relation to physical properties alone.

The unfolding of events in the film results in a scenario whereby the distinction between protagonist and antagonist is not always clear. Consequently, the events do not necessarily

follow the actions of a single individual, but rather the narrative's chain of events in which the character-focus switches at various moments throughout the film.

Act one or the establishment phase (EPSPO) depicts a 'near future' Johannesburg in which escalated levels of crime have resulted in the deployment of a robotic police force (scouts) to help maintain law and order. These robots are able to carry out instructions independently, but are only used in this law-enforcement capacity, and are clearly manifestations of weak AI.

The company behind the creation and sale of the robots, Tetravaal, is shown as becoming rapidly successful due to the scouts. Vincent, the creator of a competing and unsuccessful project within the company (called MOOSE) seems notably displeased, and distrusting of the notion of robots having the level of autonomy that the scouts have. His robot is completely controlled by a human 'pilot'. However, the police are not interested in purchasing this weapon.

In order to establish the amount of crime in the city, a criminal group is introduced. This group is comprised of three members, Ninja, Yolandi, and Amerika. It is revealed that they need to pay R20 million to a powerful gangster named Hippo. In order to do so, Yolandi suggests that they locate the creator of the scouts to find out how to turn them off, in order to successfully steal the necessary money in a heist.

Meanwhile, the creator of the scouts (Deon Wilson) has expressed interest in creating a strong AI ─ a sentient, thinking, feeling robot. Deon is shown as spending his free time working on creating artificial consciousness, which he eventually achieves. This moment brings about the first plot point (first explosion), as the creation of sentience is essentially what puts the events of the rest of narrative into motion. Deon approaches the CEO of Tetravaal, Michelle Bradley, asking whether he could test the AI on a police robot. She refuses, but Deon decides to steal a damaged scout anyway (and the 'guard key' needed to update any software within the scouts).

As Deon exits the laboratory, he is hijacked by Ninja's gang, and is taken to their hideout in a dilapidated, abandoned building. The gang discovers the damaged scout and, holding Deon at gunpoint, compels him to reprogram the robot to assist them. This begins the second act (NPSPO). The robot, now activated and sentient, is frightened of the environment and Deon explains to the gang that, firstly, it is much like a child and will need to be taught, and secondly, due to its damage, its battery cannot be replaced and it will only 'live' for five days. Yolandi names the robot Chappie, as he looks "happy like a happy chappie".

Vincent discovers the scout's activation from a Tetravaal computer, and the use of the guard key to reprogram the damaged scout. Thereafter, he asks Deon to borrow the guard key (still inside Chappie). Deon refuses, and Vincent tries to intimidate him. The now irate Deon returns to the gang's hideout, and Vincent follows him. Vincent watches in disbelief from a distance as Deon and Yolandi teach Chappie how to paint. Deon tries to convince Chappie not to conform to the gang's lifestyle and avoid crime, reminiscent of an overbearing parent. Vincent's technophobia is clear, as is the fact that this has been informed, in part, by his religious beliefs.

Annoyed by the slow progress of Chappie's learning specifically in being capable of assisting them in a future heist, Ninja takes Chappie to a group of criminals, leaving him to fend for himself. The criminals, under the impression that it is a normal police scout, take advantage of Chappie's mild mannerisms and attack him. Chappie flees. Thereafter, Vincent tracks down Chappie, removing his arm, and stealing the embedded guard chip. Chappie breaks free, and jumps out of Vincent's van, returning to the hideout. Amerika replaces Chappie's arm with a spare one left by Deon, and Yolandi comforts him. This moment signifies and highlights the difference in parenting styles between Ninja and Yolandi, with conflicting messages being sent to this newly sentient being.

Ninja asks for Chappie's forgiveness and continues to teach Chappie to work towards assisting in the heist. Chappie now acts like Ninja, and steals cars (under the impression that they belonged to Ninja) in order to raise money for explosives to be used in the heist. In order to convince Chappie that he should help with the heist (since Deon told Chappie not to engage in criminal activities), Ninja tells Chappie that Deon created him to die since his battery cannot be replaced, and that they can buy a new body if he does the heist. Chappie agrees.

Meanwhile, Vincent uses the stolen guard key to deactivate all scouts, including Chappie. This sparks a wave of crime throughout the city, as the robotic police force is now inactive. Deon takes Chappie back to the Tetravaal lab to repair him, and Vincent (hiding) sees the process. Chappie, upset about his imminent fate, steals the device used to control MOOSE in order to attempt to transfer his consciousness into a different machine. He flees back to the hideout, and tests the detection and transfer of consciousness using the gang's computer after accessing the Internet and learning the skills he needs to do so. Apart from his speed in learning language and imitative behaviours, this is the first indication of Chappie's true super-intelligent capabilities.

The gang now carries out the heist, and footage of Chappie committing crime is shown on a television news broadcast, naturally triggering outrage that one of the police scouts is assisting with a crime. This directly translates to a fear of the loss of control and the malleability of unregulated AI. Michelle gives Vincent permission to use MOOSE to destroy Chappie. This is the second plot point (second explosion), as Chappie's plan to restore his consciousness is now being prevented by a human-assisted machine.

Upon returning to their hideout, Ninja's gang is confronted by Hippo asking for his money as well as Chappie, and everyone is attacked by MOOSE shortly thereafter. The climax of the film unfolds as the gang try to simultaneously destroy MOOSE and flee the attack. While MOOSE is eventually damaged and offline, Yolandi has been killed and Deon is busy dying.

The third act (FPSPO) begins with Chappie and Deon returning to the Tetravaal lab, where Chappie kills Vincent, and uses the MOOSE control device and his new knowledge to transfer Deon's consciousness into a spare, unused scout unit. This, despite the fact that Chappie's battery is almost depleted. He essentially sacrifices himself for someone he cares about. However, the now robotic Deon transfers Chappie's consciousness to a nearby deactivated scout using the guard key.

Thereafter, Chappie returns to the hideout, and Ninja hands him a USB drive on which Yolandi's consciousness was stored while Chappie was still investigating the possibility of identifying consciousness. Chappie hacks into Tetravaal's lab remotely, creating a new robot that has a metal mould of Yolandi's face, and using the USB drive, uploads her consciousness to the newly built robot.

The conceptual relevance of the film stems from the exploration of consciousness and the nature versus nurture exploration. Essentially, in considering what it means to be human, protecting one's consciousness might result in committing acts that you may not want to, or that may not have been part of your upbringing. This emerges from Chappie's decision to commit the heist to save himself, despite repeatedly claiming that he is not comfortable with committing crime (as Deon told him this is bad).

While he goes against his own moral compass, both the moral inclination not do so, and the eventual decision to do so, are the result of socialisation from the people around him (Deon and Ninja respectively). Therefore, the audience my derive relevance through the decisions one

must take to protect themselves, and how these decisions may conflict with what we were taught as part of our upbringing.

Chappie's dialogue serves the dual function of highlighting his child-like demeanour, while also expressly stating his innate desire to live, and his consciousness. The first function emerges on Chappie's first day of activation when he repeats words such as "watch" and "mommy". This carries on throughout the film as he attempts to copy phrases such as "gangster number one" in an attempt to emulate (and gain the acceptance of) his new family.

The second function emerges through phrases like "Chappie has fears" and "I want to live. I want to stay here with mommy. I don't want to die". This highlights his self-awareness, while also strengthening the first function. Chappie also, through dialogue, provides an indication of his moral compass by asking, "Why do you humans do this? Why do you all lie?"

On the other hand, Deon's statements imply that this is the world's first strong-AI. The phrasing thereof requires careful consideration. Deon states that Chappie is the "world's first proper, full artificial intelligence". Naturally, this reinforces the myth that only once AI reaches levels of human consciousness, can it be considered as reaching the status of 'proper' or 'true'.

Deon also reinforces the nature versus nurture view taken, arguing that humans will mould the actions of sentient machines (echoing similar sentiments to Sophia, as philosophically encoded into the robot by the team at Hanson Robotics). Deon describes Chappie as "like a child in the beginning. Like a human baby. Its AI is capable of learning faster than any organic intelligence, but it'd have to be taught". This description of Chappie as explicitly child-like is useful in moments in which the audience is drawn to sympathise with Chappie. For instance, once he returns from the area in which Ninja left him, Yolandi is upset with Ninja and yells, "He's just a child".

Vincent's explicit dislike and suspicion of AI is expressed in dialogical moments such as "I have a robot that is indestructible. It is operated by a thinking, adaptable, humane, moral human being" and "your simple AI programme makes you think you're real".

The central conflict emerges as humans against a particular machine (Chappie). However, in this instance, humans are (for the most part) portrayed as bad while Chappie is portrayed as good, or pure at heart. Deon sets the entire chain of events into motion by creating sentience.

Thereafter, the sentient being's goal is to self-actualise, yet he is prevented from doing this due to the metaphoric representations of the ugly side of humanity.

Ninja represents a negative interpersonal influence, Deon represents an overbearing caregiver, Vincent represents jealousy and intolerance of the other, and Michelle represents greed as she favours her company's image over the possibility of the sentient being's survival. Yolandi and Deon both care for Chappie throughout the film, but for the most part, the humans Chappie encounters treat him poorly and actively try to prevent Chappie from survival or actualisation, in order to serve their own interests.

As a general observation, the film is depicted from an omniscient point of view, with particular character viewpoints shifting throughout. Initially, we follow events from Deon's perspective with Vincent being the force preventing him from achieving his goal. This perspective is later shared with Chappie, Ninja's gang, and Vincent all getting moments at various points wherein the events unfold as they would personally experience them. Some events are also depicted via a news broadcast, suggesting a vast detachment from a particular character's point of view.

The actual portrayal of perspective from a first-person point of view through the medium is noticeably prevalent in the film. There are a handful of shots shown from Chappie's perspective (Figure 7.1) that highlight Chappie's emphasis on facial and audio recognition. This suggests that Chappie views the world as a human would while equipping a device capable of portraying a form of augmented reality. This is notably different to a moment in which a non-sentient scout's perspective is depicted (Figure 7.2), which seems far less personal.

Figure 7.1: Chappie's point of view (*Chappie* 2015)

Figure 7.2: A non-sentient scout's point of view (*Chappie* 2015)

We also view an action from Vincent's perspective at a particular moment in the film (Figure 7.3), although this is to highlight his awareness of Chappie's existence and to portray his reaction thereto. During the climax, we also see the events from the point of view of the MOOSE (Figure 7.4), although this is reminiscent of security camera footage.

Figure 7.3: Vincent's point of view when spying on Deon (*Chappie* 2015)

Figure 7.4: Vincent's point of view when operating MOOSE (*Chappie* 2015)

The film is set in Johannesburg in the 'near future'. Considering that, during a news broadcast, the deployment of the scouts was said to have occurred "in 2016", and that a news ticker bar during the third act states that "inflation for advanced economies is projected to increase by the end of 2017", it is most likely that the events in the narrative occurred in the year 2017.

The events in the film (excluding the prologue) take place over roughly a week, since Chappie's battery only lasts for five days. The imagined future is incredibly similar to contemporary society. Besides the prevalence of robotic police (scouts) (who are explicitly named as the first of their kind), there are no other technologies, vehicles, or clothing significations that point to a difference between this imagined future and our current reality (in 2015 when the film was released, but transferable to 2020).

The events of the film take place, for the most part, in rundown areas of the city, depicting a harsh environment (Figure 7.5). This emphasises the nature versus nurture exploration, by visually emphasising the impact of one's environment on their actions and upbringing, with Chappie adapting to this harsh and unstructured environment.

Figure 7.5: Example of a crime-ridden Johannesburg in *Chappie* (2015)

7.2.2 Performance

Chappie (Sharlto Copley) is depicted as naïve, curious, caring, and child-like. He is portrayed as wanting to do what is 'right', although this is corrupted based on his interactions with humans. His true selfless character is revealed during the third act when he saves Deon by transferring his consciousness into another machine, despite the fact that his battery has almost been depleted. He is also portrayed as sensitive, arguably due to Yolandi's influence, and ashamed of his sensitivity around Ninja, while simultaneously trying to act tough to impress him.

Yolandi and Ninja (as played by themselves in accordance with their personas in the alternative hip-hop duo Die Antwoord) are portrayed as highly different throughout the film. Yolandi is caring and nurturing, adopting a motherly role, and Ninja expresses no interest in adopting a parenting role except for when it might benefit him. He is reminiscent of an abusive parent through his actions and tone.

Deon (Dev Patel) is highly intelligent and interested in creating sentience. He cares enough about Chappie that he is willing to return to the gang's hideout, despite Ninja threatening him. Vincent (Hugh Jackman), on the other hand, is portrayed as the binary opposite — highly suspicious and distrusting of autonomous AI, and even more so of Chappie's now apparent consciousness. This is clearly moulded by his jealousy of Deon's success, as well as his religious beliefs as he is seen praying once he discovers Chappie's sentience, and before operating MOOSE.

The emotional relevance of this film manifests under a broad category of universal truths and character fulfilment, in that it divulges a lesson that might be summarised as ‘you are the product of your environment’, and that ‘you might need to go against what you have been taught, to survive’. This is a relevant behavioural adaptation as it allows the audience to reflect on the influence they might have on those around them, as well as how far they’d be willing to push their own moral compass to protect their consciousness.

The police scouts in *Chappie* (2015) are depicted as a servant class, who display military-like posture and are completely at the mercy of their owners (Figure 7.6). However, they are seen as more aggressive by nature, and as having authority over the large number of criminals (mainly depicted through gangs and drug activities) in the city, and thus they are servants only to those in power (the government and police), and not people more generally.

Figure 7.6: Military posture and obedience of police scouts in *Chappie* (2015)

Chappie acts very differently to the ordinary scouts. At no point does he display the same posture and uniform voice as the rest of his robotic counterparts. He is portrayed as timid and weak during his activation, yet curious (Figure 7.7). He crawls around on the floor, and hides. This behaviour is reminiscent of something between a human child and a new (and frightened) pet. Chappie is also capable of far more expression compared to other scouts.

Figure 7.7: Chappie's frightened demeanour during his activation (*Chappie* 2015)

Chappie's demeanour, posture, and accent change to emulate Ninja towards the end of the film, further highlighting the effect of one's environment on their development. This manifests in Chappie acting in a way that Ninja has told him is 'like a gangster' (Figure 7.8), including posture, slang, and vocal intonation.

Figure 7.8: Ninja teaching Chappie how to act like a 'gangster' (*Chappie* 2015)

Chappie's character arc emulates the development of a child, from the point of view that a child would grow up to become the product of their environment, while still instinctively fighting for survival. In his case, he transforms from curious and naïve, to ultimately a selfless saviour for both Yolandi and Deon despite clearly displaying the will to survive throughout. He also reaches levels of superintelligence, to a point of being able to usher in the beginning of a transhumanist singularity.

On the other hand, Vincent (ultimately Chappie's biggest threat to survival) has a consistent arc, motivated by revenge and phobia of the other. Instead of the arc changing due to personal development, Vincent's arc moves towards even greater intolerance of the other and self-centeredness.

7.2.3 Medium

For the most part, the semiotic resources as part of the medium of the film mainly support the progression of the story as their primary function, with little additional meaning divulged throughout. However, shots during key moments often depict, rather strongly, relationships between characters.

For instance, Figure 7.8 and Figure 7.9 highlight the different 'parenting' styles of Ninja and Yolandi. The framing of Figure 7.8 depicts Ninja leading Chappie towards (perhaps metaphorically) a particular direction, and Figure 7.9 depicts Yolandi as a nurturer and literally (and metaphorically) close to him.

Figure 7.9: Yolandi comforting Chappie once he returns from Vincent's van (*Chappie* 2015)

The cinematography also works to highlight Chappie's innocence during his activation, with high angles depicting his vulnerability, showing him from the perspective of how one might view a child or a small pet, despite his adult-like size in actuality (Figure 7.10). This shifts to a low-angle portrayal of the robot during the resolution, as Chappie seeks revenge against Vincent (Figure 7.11).

Figure 7.10: High angle portraying Chappie's innocence (*Chappie* 2015)

Figure 7.11: Low angle portraying Chappie's power as he seeks revenge (*Chappie* 2015)

Chappie is CGI based on the performance of Sharlto Copley (Figure 7.12). Visual effects were used in the design of Chappie's point of view, as well as the MOOSE camera perspective. Notably, the visual design used in Chappie's perspective has additional details such as key words learnt and relevant in the context (Figure 7.13, on the bottom left). This progresses as Chappie learns throughout the film. This is also infused with additional information that one might be able to access on a desktop computer, such as CPU usage.

Figure 7.12: Behind the scenes footage of Copley as Chappie (FilmIsNow 2016)

Figure 7.13: Visual effects used for Chappie's point of view (*Chappie* 2015)

The sound design of the film supplements the harsh visual atmosphere, while also assisting in distinguishing characters from one another. There are snippets from Die Antwoord's music throughout the film. This, of course, makes sense since the Zef alternative hip-hop duo are central characters in the film's narrative. Their music also helps to establish a particular tone, further emphasising the harsh environment in which the innocent and impressionable sentient robot finds himself. This also helps locate the film within a uniquely South African environment, as Zef culture is unique to the country. However, other pieces of music are also used throughout to establish the mood at particular moments, especially during action sequences and sombre moments.

Apart from the music, the voice of the robots is a notable sonic signifier. The entire film is based on the premise that a being exists beyond their physical body. Chappie's voice

(frequency, pitch, and tone) remains the same once he is transferred into the body of a different scout model. This makes sense considering that he was originally an ordinary scout himself, so the technology that would have been used as (essentially) a speaker would remain the same and be capable of maintaining the sonic codes denoting his voice. This becomes more complex, however, when Deon is transferred into a scout's body. Deon's voice remains the same, although reverb is added to make it sound more 'robotic'. From a physiological point of view, it could be argued that Deon's voice (excluding accent) should be more similar to the voice of Chappie. However, this allows the filmmaker to represent the scout housing Deon's essence as housing the *true* Deon.

Due to the complexity of the story in terms of the journey of both Ninja's gang and Deon's quest to teach Chappie right from wrong, there is quite a large amount of parallel editing between scenes as they unfold at the same time. For instance, just before the first plot point, tension is created as we see cross cutting between Deon's plans, and the imminent prevention of his goal due to Ninja's gang's simultaneous planning. Thereafter, the different characters (living in their seemingly different worlds) are brought together. This happens again during the second plot point and climax, as Deon simultaneously tries to reactivate the scouts while Ninja's gang carries out the heist. Deon rushes to Ninja's hideout after seeing the heist on television, and this is what allows Vincent permission to use MOOSE.

The particular stylistic decision to incorporate fictional news inserts throughout the narrative and documentary style interviews in the beginning helps establish a perceived realism (as a result of inter-medium intertextuality), suggesting that the events within the narrative are plausible. For instance, Figure 7.14 depicts a moment in which Anderson Cooper (an actual CNN television journalist) discusses Tetravaal's success in a fictional CNN broadcast. This is also a stylistic trademark of Neill Blomkamp (the director), who used a similar mockumentary style in the film *District 9* (2009).

Figure 7.14: Anderson Cooper's fictional CNN broadcast in *Chappie* (2015)

7.2.4 Aesthetics

Colour is used to establish particular moods in certain moments throughout the film. For instance, the colder blue colouring inside Vincent's van is distinguished from the warmer orange and green tint of Yolandi's room after Chappie's returns from Vincent's first attack (Figure 7.15). This connotes the warmth of a home, even when such a home is dysfunctional in and of itself.

Figure 7.15: Colder tones in Vincent's van (left) juxtaposed with warmer tones in Yolandi's room (right) in *Chappie* (2015)

However, throughout the film, mainly in exterior shots in the city, the colours are often heavily desaturated. This highlights the harshness of the environment as dull and ominous, and this works to further emphasise the impact thereof on a child's development.

While adopting a humanoid body shape, the design of the scouts is rather distinguishable from humans in that their facial features differ (Figure 7.16). They do not have a mouth, nose, eyes, or ears (in a human sense). Instead, their eyes are replaced with a single screen with dots used to indexically denote vision. A mechanical bar serves as a brow, while the same type of bar

and an additional screen serves as a mouth on the opposite end of the face. Furthermore, the robots were designed with what appears to be antennas on the top of their heads.

Figure 7.16: Regular police scout in *Chappie* (2015)

Chappie is distinguishable from other scouts in that his eyes are not depicted as a single beam, but as square blocks (Figure 7.17). This denotes that he is indeed more similar to humans (in relation to other scouts), allowing humans to peer into what might indeed be more akin to human eyes for the purposes of empathising with the machine.

His ears also change direction based on his emotion (perhaps metaphorically similar to a fearful animal, particularly a rabbit), and he is capable of expressing emotion. He also has a sticker on his head which reads 'REJECT', since he was scheduled for decommission, which ties in heavily with him being an other, with whom Vincent (representative of an in-group) lacks tolerance. This is further supported by him describing himself, metonymically, as a black sheep at the end of the film.

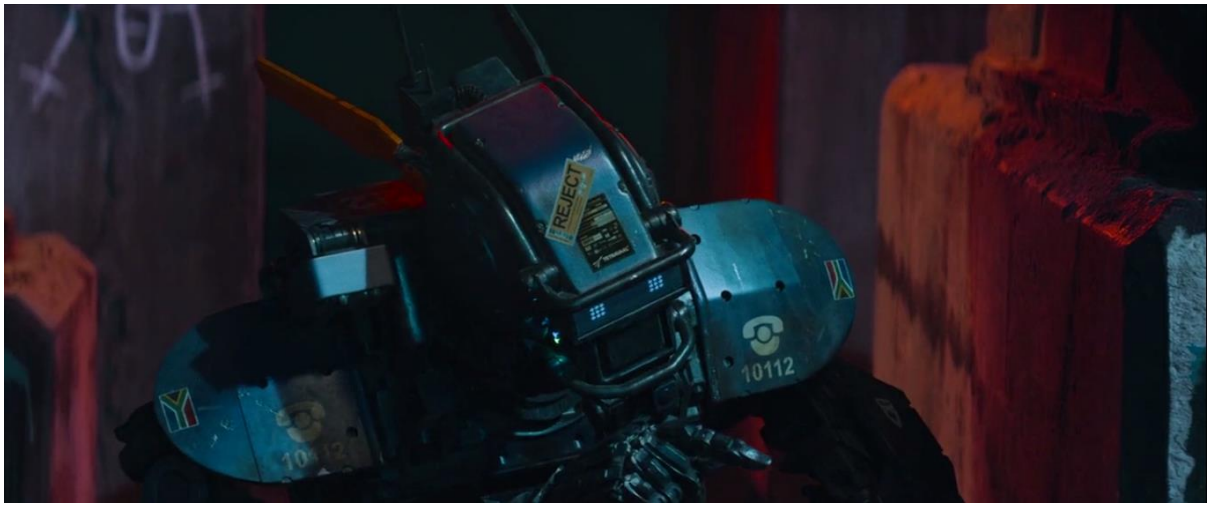


Figure 7.17: Chappie is distinguishable from other scouts (*Chappie* 2015)

Ultimately, once socialised by Ninja's gang, Chappie undergoes a further stylistic transformation as he is given jewellery (or 'bling', as referred to in the film), and Ninja spray-paints symbols denoting Zef culture and gang culture on Chappie's exterior (Figure 7.18). This connotes his mental transformation as he is socialised by Ninja. He also adopts a mismatched arm as the narrative progresses, after Vincent removes it.

Figure 7.18: Chappie's transformed appearance as he is socialised by Ninja (*Chappie* 2015)

Yolandi's robotic design (created by Chappie) includes an artificial face modelled according to Yolandi's human face. This includes eyeballs, a nose, and even lips (Figure 7.19). However, the entire face is a single shade of white. This design is remarkably different to any of the other robots in the film. This is notable, as Yolandi's activation is the very last shot of the film before the credits begin rolling. This suggests the possibility of a mass future singularity moment, or

even perhaps serves to blur the distinction between human and machine even further by emphasising the mind/body split.

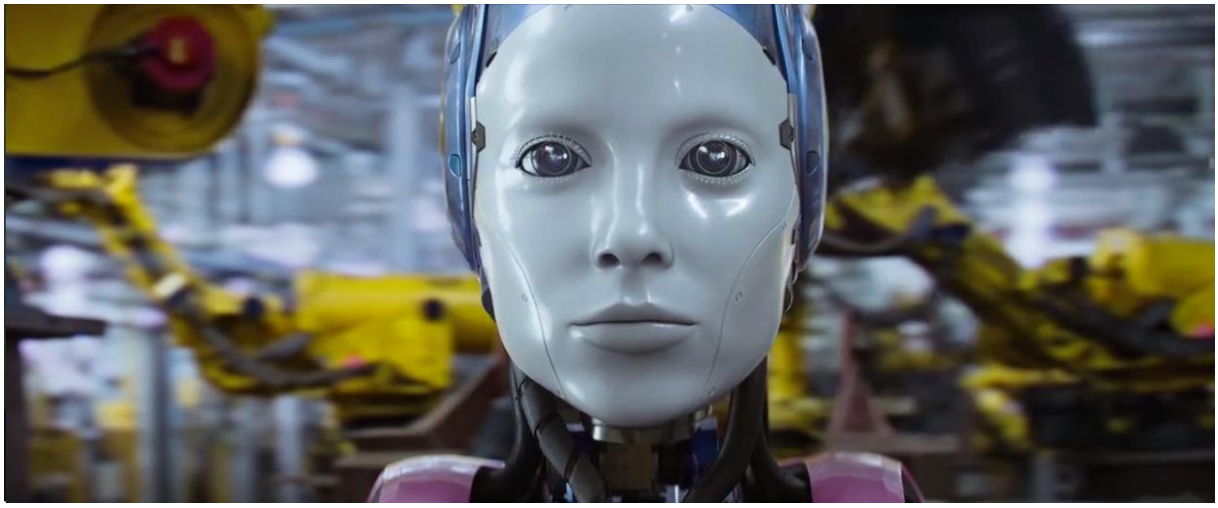


Figure 7.19: Yolandi's robotic face (*Chappie* 2015)

A central hero prop within the film is the 'neurotransmitter', which Chappie successfully uses to transfer consciousness for himself, Deon, and Yolandi. Since the diegesis only portrays the introduction of police scouts as differing from our current reality technologically, most of the other props either serve as functional narrative tools, or denotations for Zef culture (as is the case with Ninja's gang's weapons).

The use of Zef imagery throughout Ninja's gang's hideout further enhances the alternative, dysfunctional environment in which Chappie finds himself. This emerges in the form of bright colours against a dull scenery, graffiti, and even the decision to create such an elaborate and intimate home setup in such a large, abandoned building (Figure 7.20).

Figure 7.20: Ninja's hideout (*Chappie* 2015)

The various shots of locations around Johannesburg clearly show poverty and hardship, when juxtaposed with the more affluent area in which Chappie steals cars for Ninja. This highlights the different worlds that one might indeed find themselves within a single city, in order to further highlight the nature versus nurture thematic considerations.

7.2.5 External factors

Chappie (2015) was released in March of 2015. As such, the level of technological advancement during its release was far higher than during the release of *I, Robot* (2004). This is due to the rapid technological advancement of the fourth industrial revolution. However, it can be argued that this also brought with it greater familiarity and acceptance of personal technologies.

Naturally, mobile technology had made significant strides since 2004. While during this period phone cameras were only beginning to emerge as a novel concept, 2014 saw the introduction of one of the first 4K mobile video capturing options. Touch-screen inputs were now commonplace and cellular phones had moved away from the race towards compactness in the early 2000's towards larger, clearer displays for HD video. According to Larson (2014), this period also saw virtual reality beginning to emerge as a mainstream concept, Google had created a contact lens that could monitor blood-sugar levels, and self-driving cars were rapidly advancing. Robotics and AI were also far more advanced during this time and Sophia was only one year away from activation. Furthermore, 4G mobile networks also allowed people to interact with each other online with far greater speed and efficacy, and social media was widely used, to a point that, in some regions for younger people, one might be considered an outsider for not being connected to certain social media platforms.

Stylistically, the film is highly intertextual through its incorporation of Zef elements that would not be out of place in one of Die Antwoord's music videos. Many of the staple science fiction intertextual references in *I, Robot* (2004) were also evident in *Chappie* (2015), including themes of Frankensteinian creation and the use of slave-robots from *R.U.R.* (1920), yet *Chappie* (2015) breaks the convention of robots turning against their makers as Chappie is depicted as generally caring about those around him. The machine used to transfer consciousness is also highly reminiscent of the transfer of Maria's face to a robot in *Metropolis* (1927), although this is taken further in the film, as the being's entire essence is transferred.

According to Blomkamp, the film is about "the birth of a pure soul or spirit", and even if robots do attempt to enslave humanity, he would view it as a step in the right direction (Truitt 2015: para. 12). He claims to not have anything invested in humans, but rather "in the idea of evolution continuing its spiral of complexity" (Truitt 2015: para. 12). Blomkamp has also stated that he believes the film is not necessarily about AI, but the meaning of consciousness, and that it explores the notion of whether a being with a different consciousness to the consciousness encased in a human shell should be deemed to have any less importance (Lambie 2017).

Blomkamp has expressed doubts as to whether strong AI, as depicted in contemporary cinema, would indeed ever be possible (Zabarah 2015). He has, however, expressed a belief that the weak AI scouts in *Chappie* (2015) might indeed become a reality (Zabarah 2015: para. 10), stating that:

> I definitely think it'll get there. That's not even a debate. It will get there within a decade or less. Like if you take Petman or [other robots] from Boston Dynamics and look at what they're doing, you mix that with some sort of complex code that has a bunch of protocols about how to react to certain situations. We will absolutely make that. That's scarier to me, weirdly, than real AI. That actually bothers me more.

Furthermore, Blomkamp has stated that the decision to portray Chappie as good and many of the humans as bad was intentional, stating that "what I really like about Chappie is that it's doing the opposite of what Hollywood films normally do. The Hollywood thing is that robots always blow us up... So I thought: 'Nah, I'll just go the other way'" (Hawkes 2015: para. 12). Finally, Blomkamp also claims that the decision to change Chappie's eyes to squares from beams was intentionally done to 'soften' him (Hawkes 2015), in order to allow the audience greater emotional access to the character.

This film is relevant to contemporary society in that it explores the age-old theme of survival and the tendency for humans to engage in othering, while transferring this to a contemporary society in which nanotechnology, conscious humanoid robots, and technological convergence are all in their early stages. This is achieved through exploring the notion of consciousness through AI, clearly a topic which is interesting to many people due to its contemporary advancements, while people become more reliant on AI to perform personal and commercial activities.

7.3 Survey responses

Eighty-three (83) people responded to the survey for *Chappie* (2015), with 55 responses considered for this study since the remainder were incomplete.

7.3.1 Demographic information

The demographic composition of respondents was, in all categories, heavily skewed towards technological optimism (Addendum D). Accordingly, 41% of respondents were under 30, while only 27% were over 40. However, as with the previous set of data, many individuals that were over 40 still found themselves within the A_{PO} as they still met the criteria of having 66% compatibility with the other indicators.

For instance, 91% of respondents were male and classified themselves as non-religious Furthermore, 80% had completed some form of tertiary education, with 84% from first world countries. Every respondent selected at least one of the options in question 1.9, confirming that 100% of respondents interact with some form of contemporary AI service or application.

From this demographic information, the vast majority (85%) of respondents were identified to be predisposed to optimism (A_{PO}), while only 5% were predisposed to pessimism (A_{PP}) (Figure 7.21).

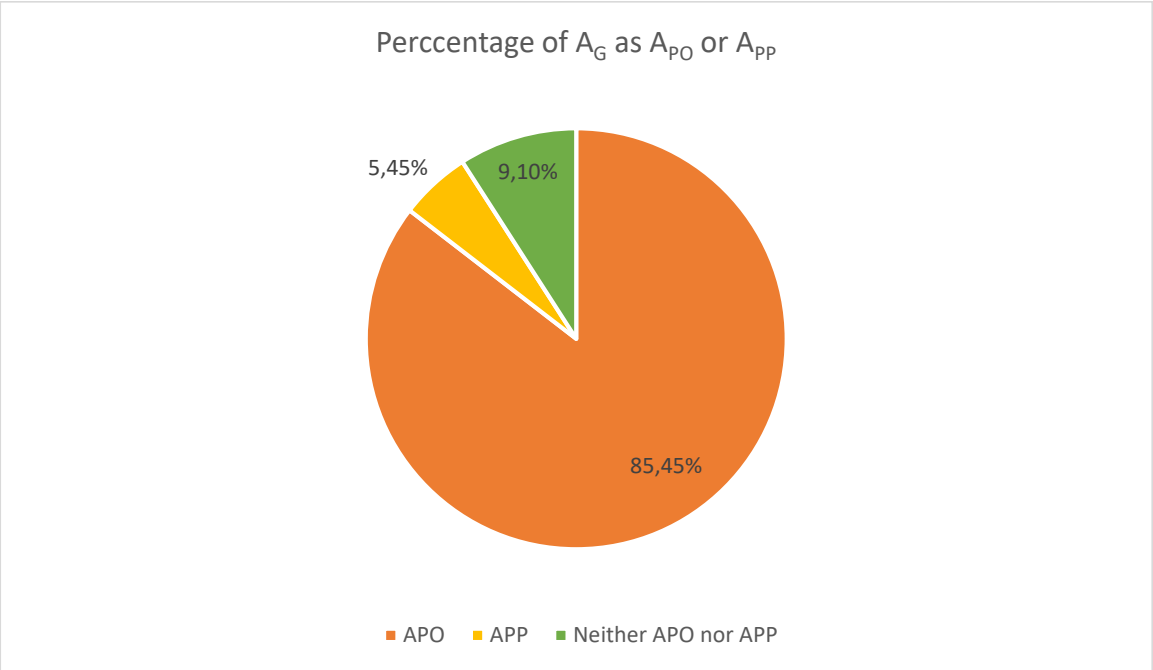


Figure 7.21: Composition of *Chappie* (2015) A_G as A_{PO} and A_{PP}

7.3.2 Perceptions of AI stemming from the film

The vast majority of respondents (78%) either agreed or strongly agreed that the film was engaging and/or entertaining, while only 16% disagreed or strongly disagreed with this statement (Figure 7.22).

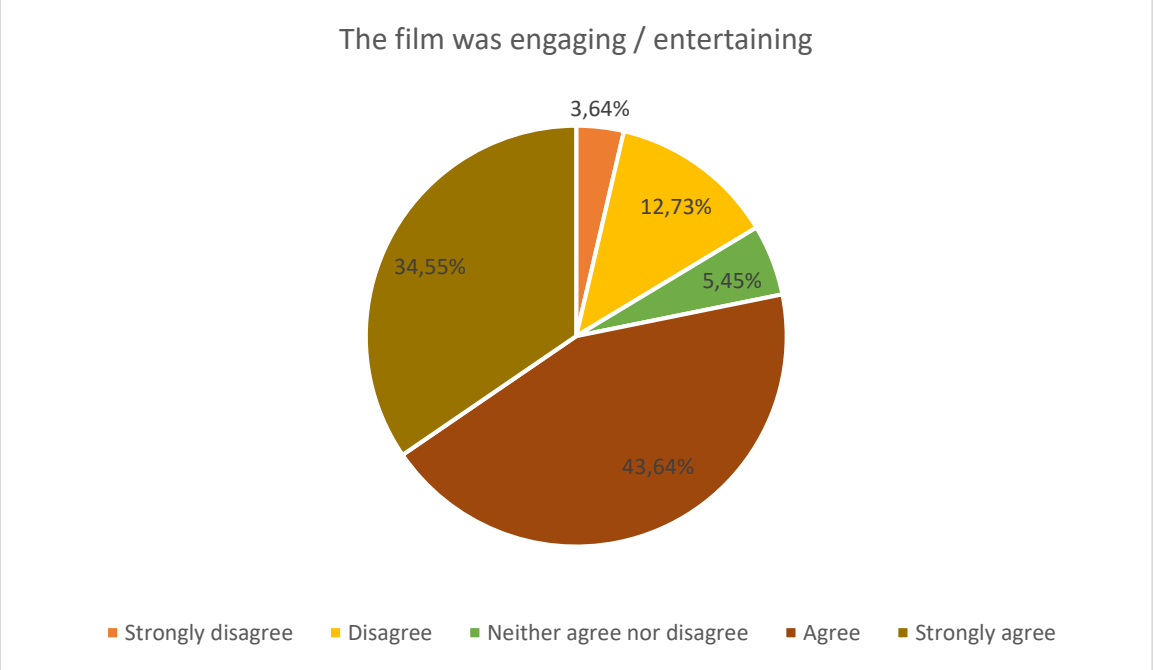


Figure 7.22: *Chappie* (2015) survey responses for question 2.1

This points towards an effective suspension of disbelief and narrative transport of the vast majority of respondents. This is confirmed by some of the qualitative responses in Addendum E, but concerns were raised about the decision to cast Die Antwoord in a film exploring this type of topic (although some support for the decision was also expressed).

Almost 70% of respondents believed that the humanoid robots in the film were portrayed positively (Figure 7.23), and 56% believed that they were portrayed more positively than most other science fiction films that they have viewed (Figure 7.24).

Inversely, only 13% did not agree (with only around 2% strongly disagreeing) that robots were portrayed positively, and only 9% believed the portrayal to be less positive than other films used as an individual frame of reference.

Responses (Addendum E) such as “Everything he thought or did was taught. So if we teach AI to be good they will be” and “I think they showed the good and the bad sides very well. How

Chappie helps commiting [*sic*] crimes because he doesnt [*sic*] know better, but also saving his maker and protecting those which he cares about" suggests that the audience might have mixed feelings about Chappie due to his criminal activities, but that ultimately it is his capacity to care that deems his portrayal as positive.

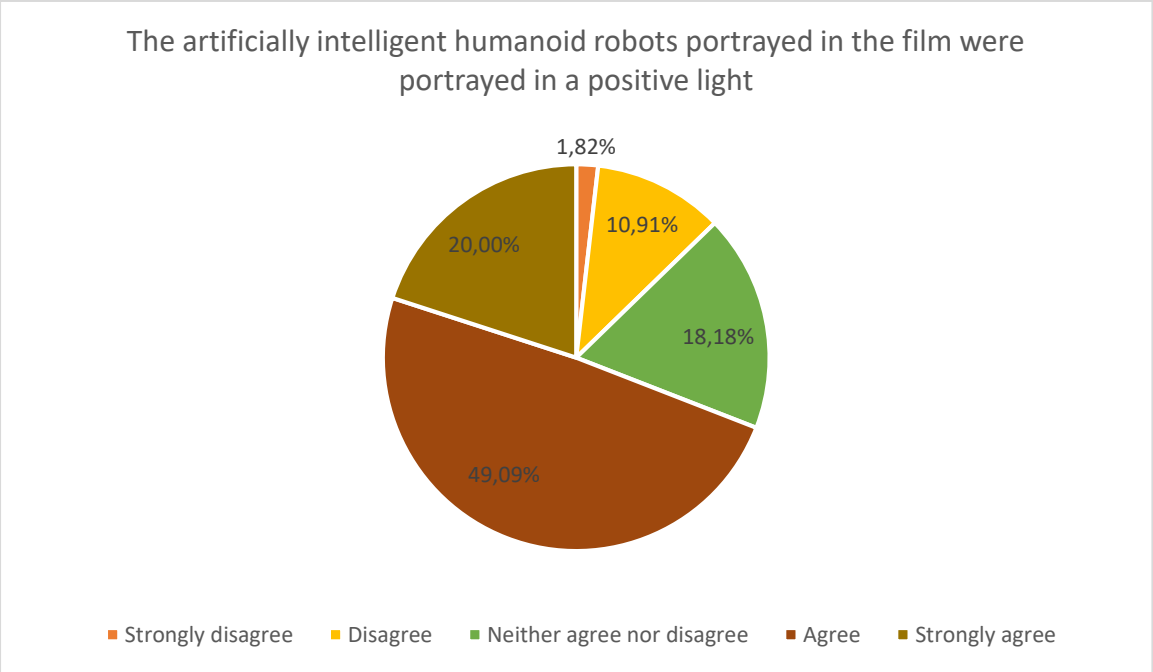


Figure 7.23: *Chappie* (2015) survey responses for question 2.2

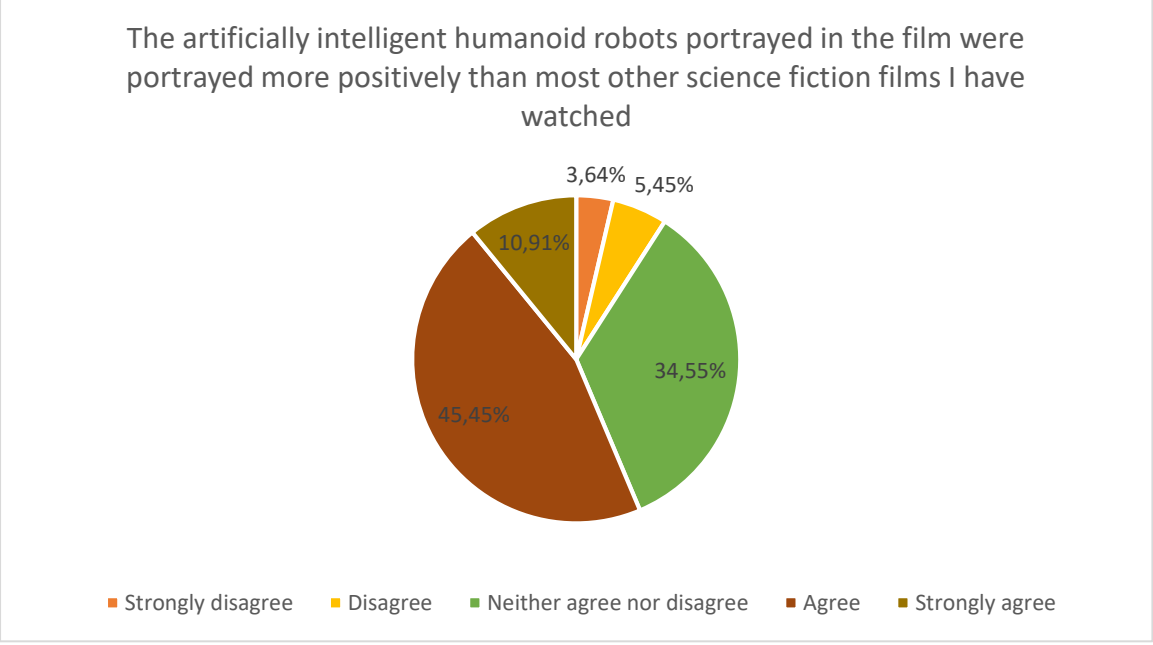


Figure 7.24: *Chappie* (2015) survey responses for question 2.3

Over 75% of respondents indicated that they cared about one or more robots in the film (Figure 7.25), with 36% strongly agreeing with this statement, and each of the qualitative responses (Addendum E) indicating that this robot is Chappie.

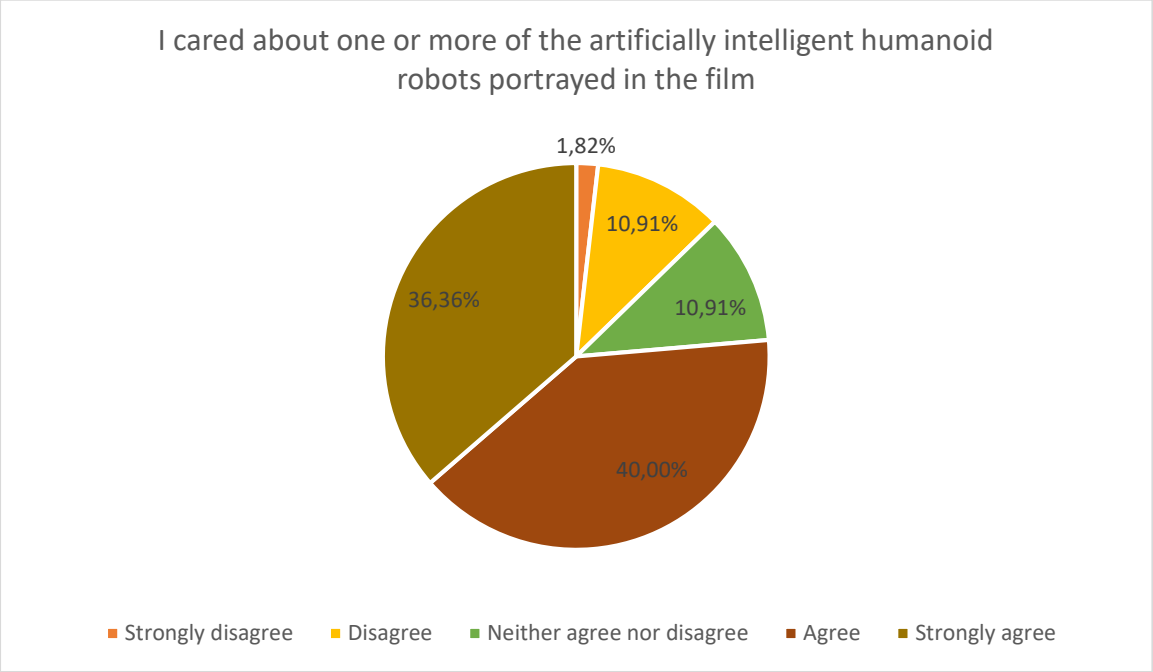


Figure 7.25: *Chappie* (2015) survey responses for question 2.4

Roughly half (51%) of respondents disagreed or strongly disagreed that the film helped increase their understanding about AI and consciousness (Figure 7.26), with 20% agreeing, and none strongly agreeing.

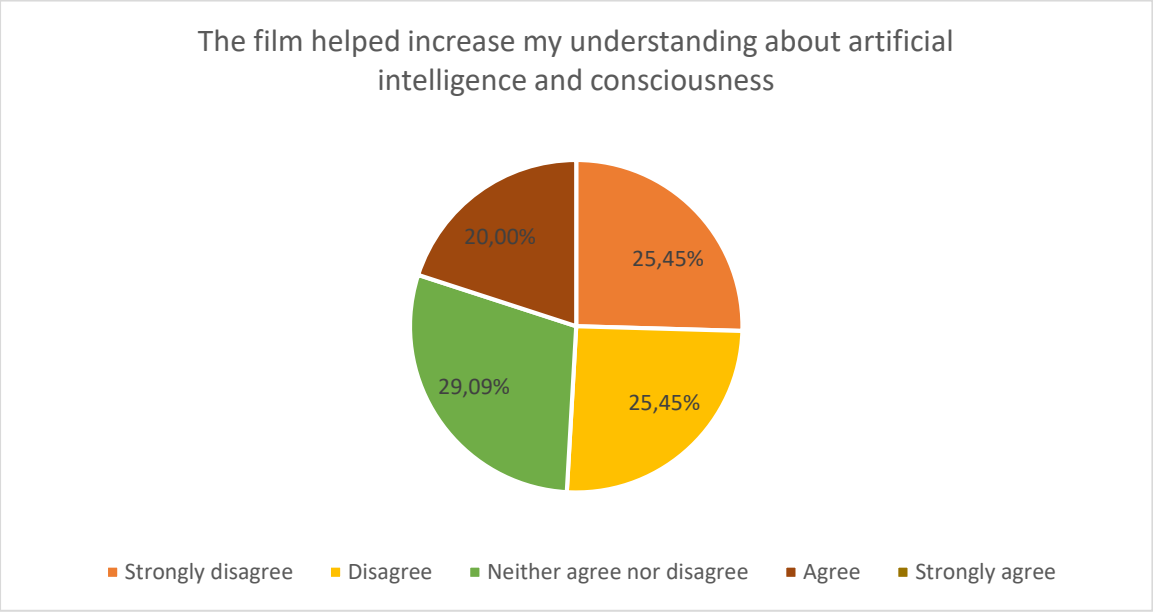


Figure 7.26: *Chappie* (2015) survey responses for question 2.5

Over a third (36%) of respondents felt less positive about future AI after viewing the film (Figure 7.27), with only 15% feeling more positive, and half (49%) feeling neither more nor less positive. The qualitative responses (Addendum E) seem to suggest that this might be because, while Chappie is shown as innocent, he is also shown as naïve and this suggests strong AI's potential to be abused by humans and consequently cause abuse. Almost 3/5 (58%) of respondents have, at least once, thought about the film when seeing, hearing, or saying something else about artificial intelligence (Figure 7.28), while only 35% have never done so.

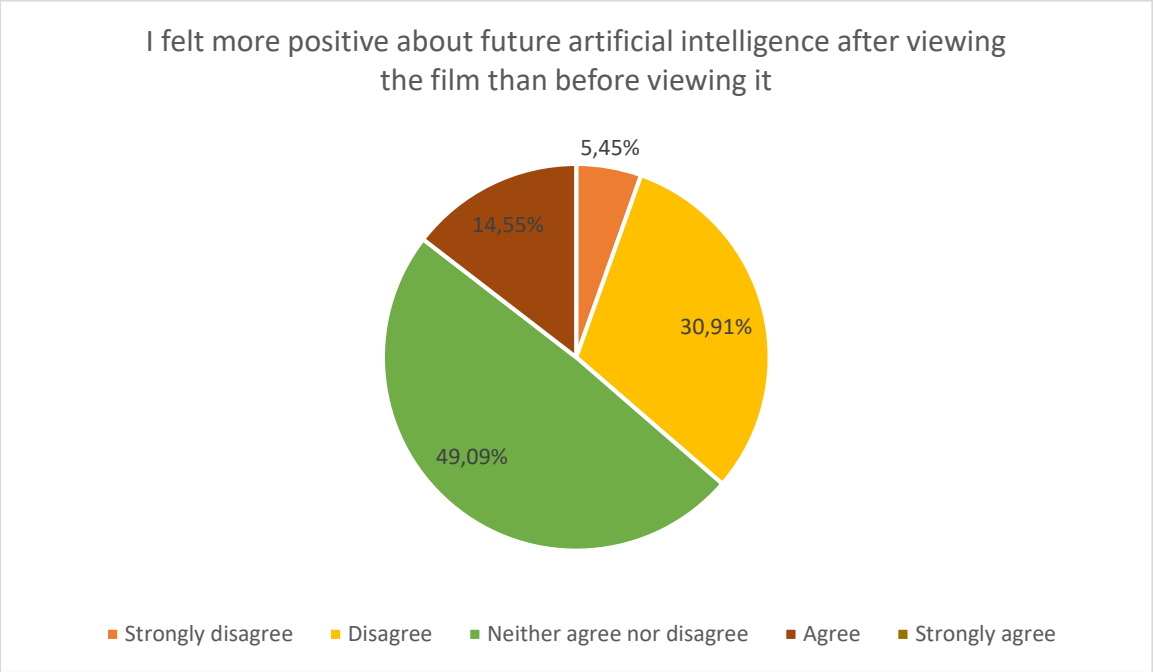


Figure 7.27: *Chappie* (2015) survey responses for question 2.6

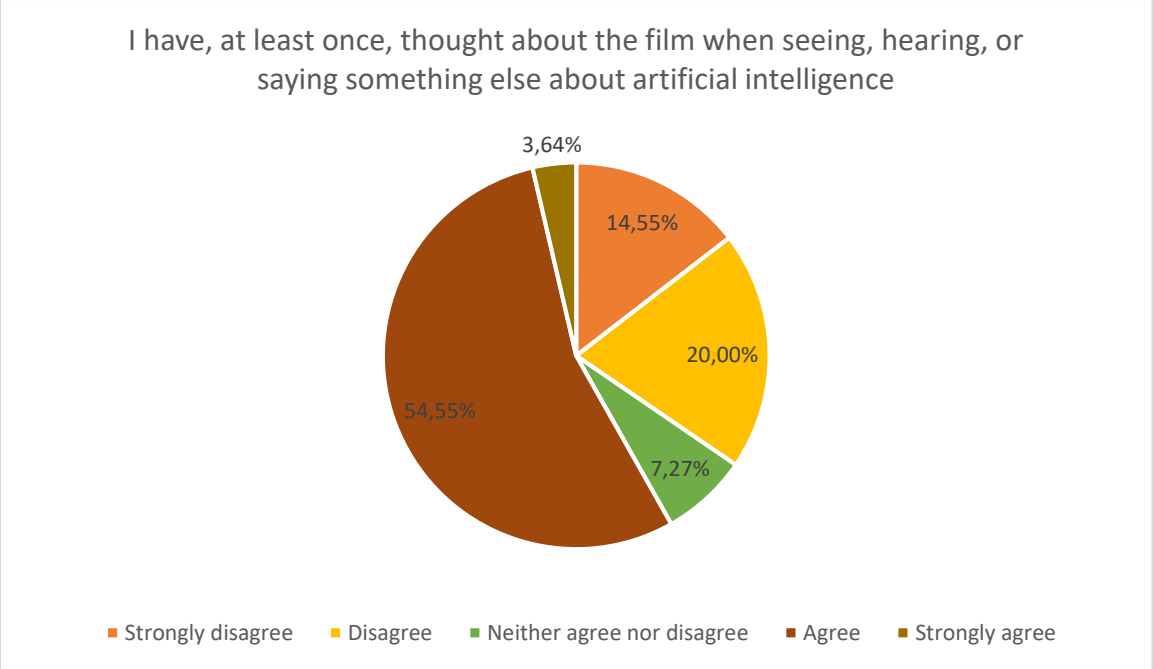


Figure 7.28: *Chappie* (2015) survey responses for question 2.7

Almost a third (31%) of respondents believed that artificially intelligent humanoid robots that behave in the same way as the robots portrayed in the film will be created (Figure 7.29). This is quite balanced considering that 31% were undecided and 38% did not believe this would be the case.

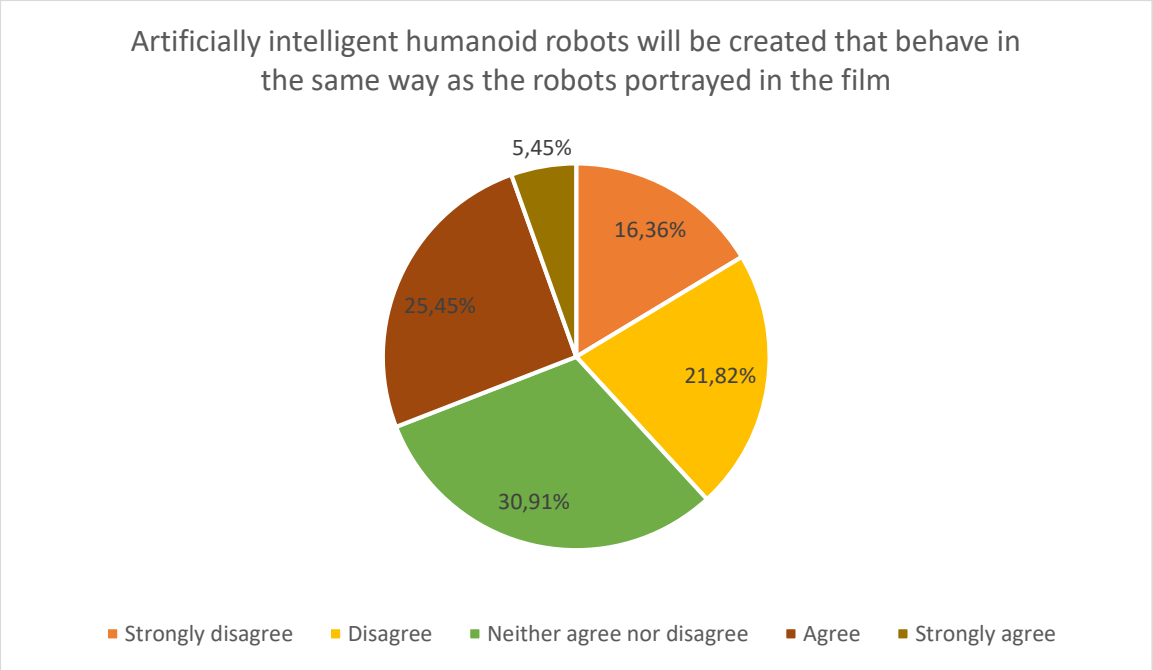


Figure 7.29: *Chappie* (2015) survey responses for question 2.8

There was further division of opinion with 41% of respondents wanting to live in a world where robots are as advanced as those portrayed in the film, and 40% expressing that this would be undesirable (Figure 7.30).

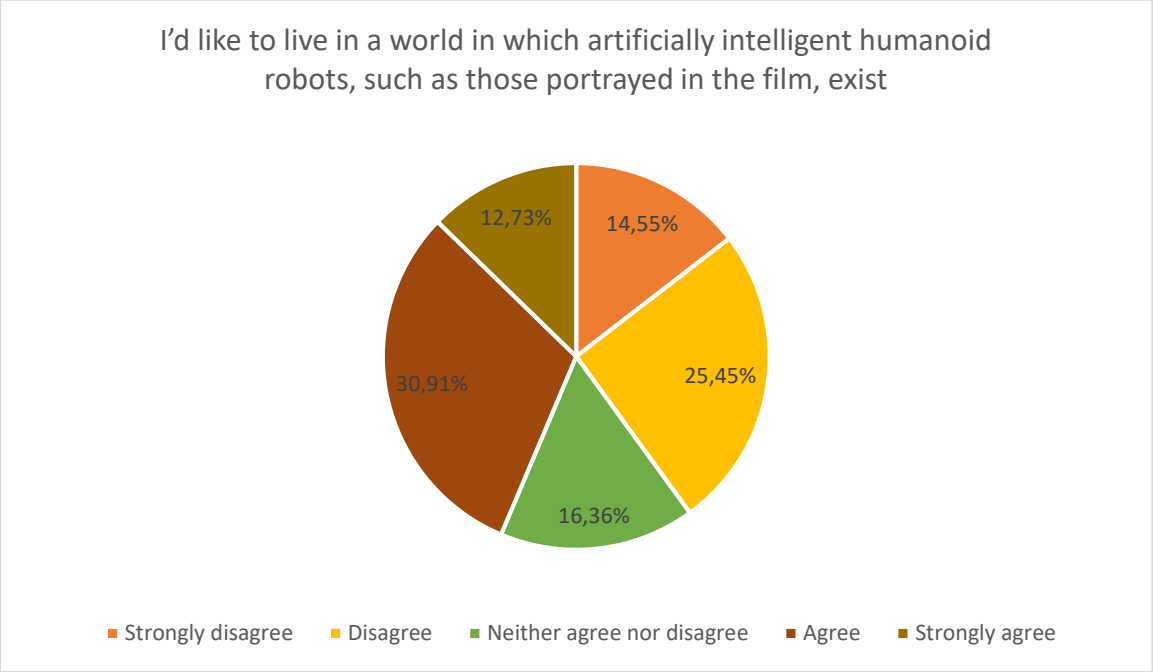


Figure 7.30: *Chappie* (2015) survey responses for question 2.9

However, only 7% of respondents claimed that the film changed their opinion of the possibility of conscious artificially intelligent humanoid robots being created (not a single respondent strongly agreed); while 56% stated that it did not change their opinion in this regard (Figure 7.31).

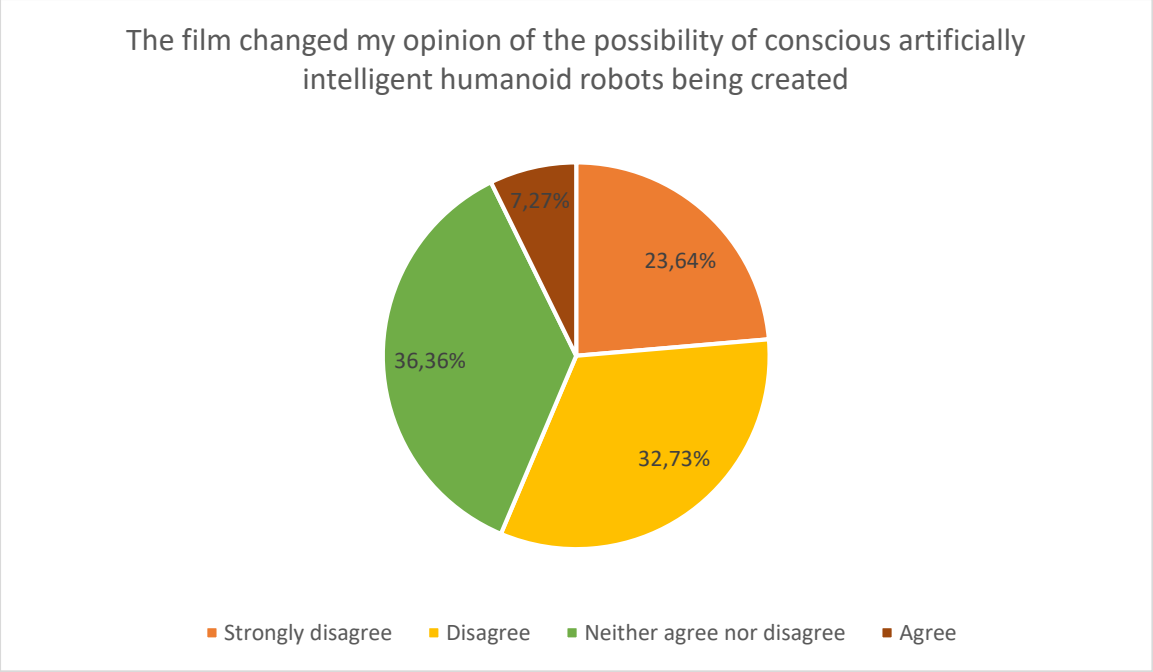


Figure 7.31: *Chappie* (2015) survey responses for question 2.10

7.3.3 Perception/knowledge of AI generally

Only 24% of respondents strongly agreed that they often interact with services or applications which make use of artificial intelligence (Figure 7.32), despite 100% responding to at least one service/application in question 1.9. However, 53% agreed with this, and only 13% disagreed or strongly disagreed.

Regardless, 85% of respondents either agreed or strongly agreed that they know about the trends and use of current and upcoming technology (Figure 7.33). Only 2% disagreed, with no respondents strongly disagreeing.

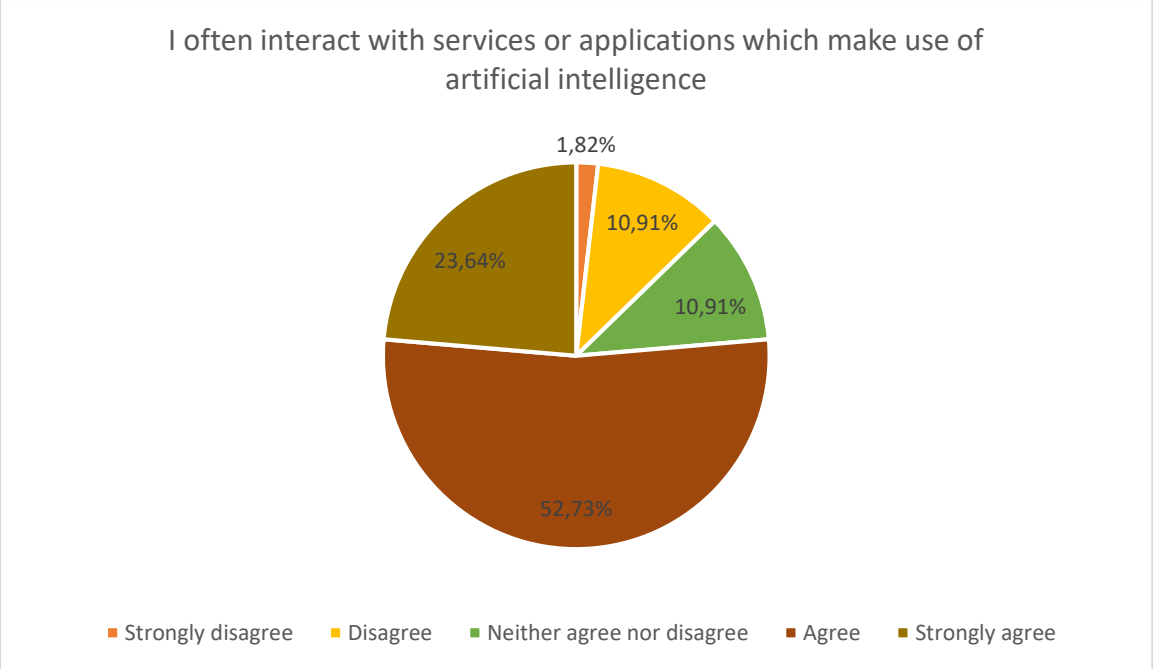


Figure 7.32: *Chappie* (2015) survey responses for question 3.1

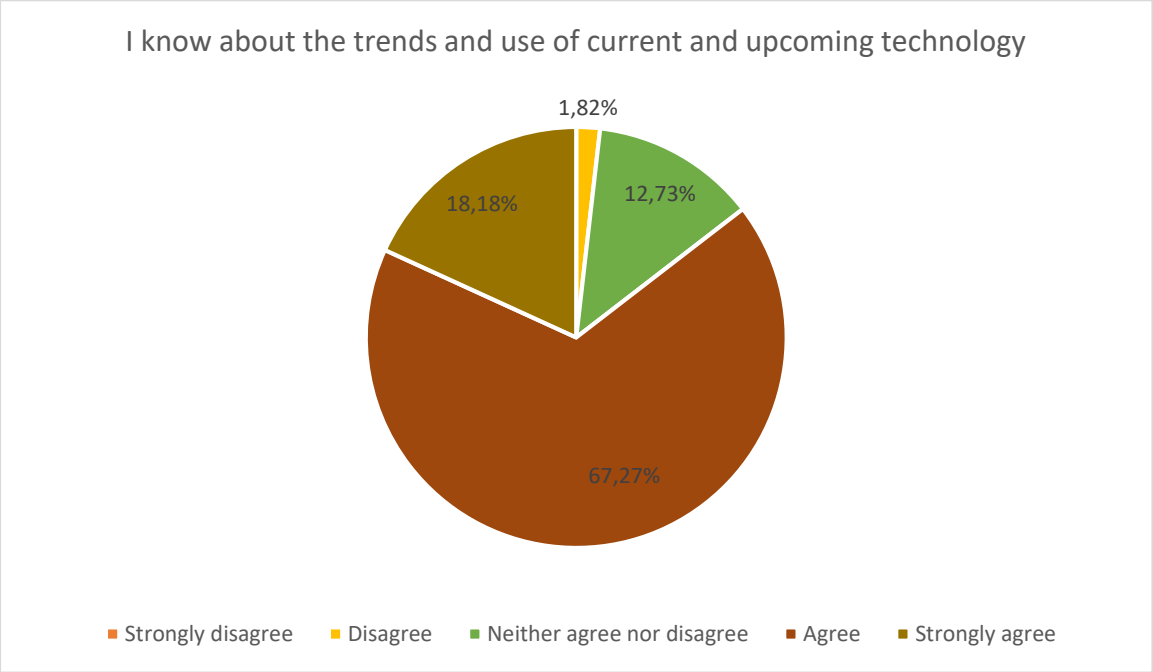


Figure 7.33: *Chappie* (2015) survey responses for question 3.2

Finally, 73% of respondents agreed or strongly agreed that they were optimistic about future technologies involving artificial intelligence, with only 11% disagreeing (Figure 7.34). Those disagreeing indicated that this is due to the potential for misuse by powerful governmental or corporate bodies. Similarly, 73% of respondents believed that AI would ultimately be good or great for humans and/or the planet, with only 16% disagreeing (Figure 7.35).

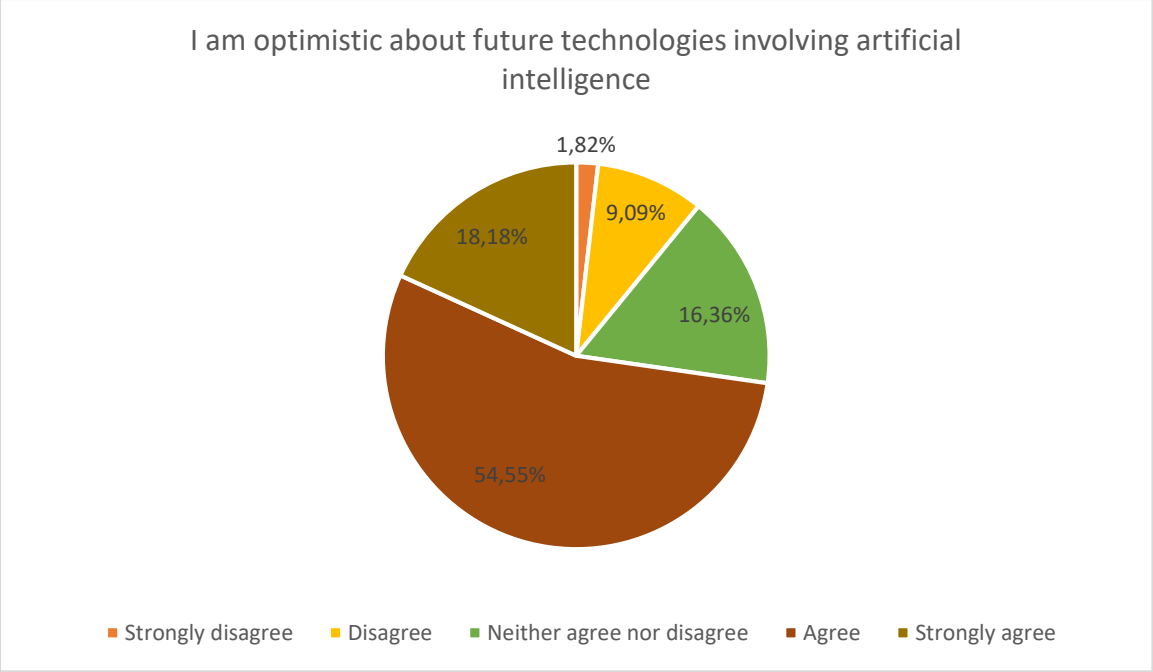


Figure 7.34: *Chappie* (2015) survey responses for question 3.3

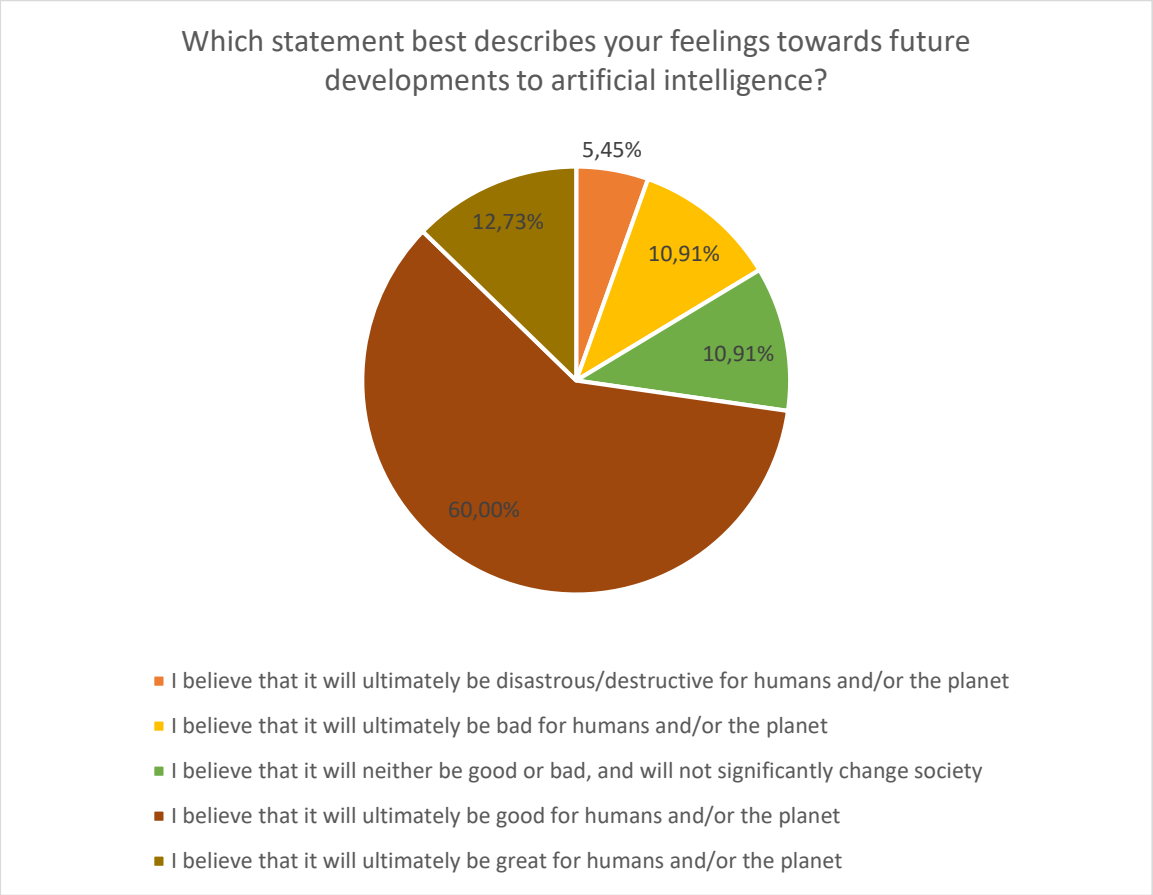


Figure 7.35: *Chappie* (2015) survey responses for question 3.4

7.4 Conclusion

This chapter served the same function as the previous chapter, in that it provided the analysis for one of the texts being analysed. Accordingly, the findings for the social semiotic analysis and survey responses for *Chappie* (2015) were presented.

In terms of the social semiotic analysis, the narrative elements served to explore the notion of consciousness and the essence of being. This exploration emerges as a strong argument for the role of environment in shaping behaviour. Instead of a regulatory framework, the AI in the film learns right from wrong by interacting with other humans. It also explores the transhumanist idea of transferring consciousness into a robotic body, allowing humans to overcome death.

These elements relate to the theme of losing control (since the AI is moulded by his environment and manipulated in various ways), yet the AI is ultimately portrayed as a hero figure that essentially wants to do whatever he understands 'good' to be, learning care naturally. This emerges through a plot in which an AI battles against humans to survive (which is also at the core of the conflict of the film), with Chappie being portrayed as the sympathetic hero and, most often, humans portrayed as iterations of the ugly side of humanity.

These themes and elements of conflict speak directly to performance elements of the film, as Chappie is portrayed as childlike, acting differently throughout the film as he learns certain behaviours through socialisation. Chappie is depicted as strong AI, and is the only robot that can be considered in this class. Of course, this excludes the former humans, now in robotic bodies, at the end of the film.

The use of the cinematic medium serves to enhance these meanings, and amplify the attempt to make the audience sympathise with the robotic protagonist. Cinematography highlights proximity (and the relationship) of the robot to those around him. Notably, the use of sound argues that one's essence can be transferred to a different body, while the use of mockumentary elements heightens the realism of this transhumanist notion being plausible in a near future.

Aesthetically, many elements are used to highlight the harshness of the environment in which Chappie finds himself, further emphasising the importance of socialisation as part of essence. The police scouts, of which Chappie is one (by form), are quite distinguishable from humans in that they have non-human facial features. For instance, instead of eyes, single beams of light denote their vision. Chappie, however, has square blocks of light ─ more similar to what one

might recognise as human eyes. Furthermore, while Chappie lacks other elements that are normally associated with humanity, he is capable of expressing emotion. This is primarily achieved through the position of his ears, which are reminiscent of an animal, perhaps a pet rabbit, further portraying innocence and non-culpability for many of his wrongdoings.

The socio-cultural context of production was marked by acceptance of personal technologies including mobile phones and social media. Intertextually, the film follows many science fiction generic staples, and the decision to cast hip-hop duo Die Antwoord means that, stylistically, the film has many elements of Zef culture. The director of the film saw it as exploring the meaning of consciousness and the purity of a new spirit. Its relevance can be explained as an exploration of the age-old theme of survival and the tendency for humans to engage in othering, while transferring this to a contemporary society in which nanotechnology, conscious humanoid robots, and technological convergence are all in their early stages.

Concerning the findings of the audience research, the demographic data indicated that the vast majority of the respondents (85%) were predisposed to optimism towards the technology, while only 5% were predisposed to pessimism. The indicators used to measure this were shown to be accurate, since 73% of respondents agreed or strongly agreed that they were optimistic about future technologies involving artificial intelligence, and believed that AI will ultimately be good or great for humans and/or the planet.

Around 70% of respondents believed that the humanoid robots were portrayed in a positive light overall. Only 9% believed the portrayal to be less positive than other films used as an individual frame of reference. Furthermore, the majority of respondents (70%) cared about one or more robots in the film, which was qualitatively confirmed to be most likely Chappie. However, only 20% of respondents agreed that the film helped increase their understanding of AI and consciousness, and 36% felt less positive about future AI after viewing the film. Almost 60% of respondents have, at least once, thought about the film when seeing, hearing, or saying something about AI, while only 35% have never done so.

Almost a third of respondents believed that artificially intelligent humanoid robots will be created that behave in the same way as the robots portrayed in the film, but only 7% of respondents claimed that the film changed their opinion of the possibility of this (not a single respondent strongly agreed). There was an almost 50/50 split as to whether respondents would

like to live in a world in which artificially intelligent humanoid robots, such as those portrayed in the film, exist.

The following chapter provides an analysis and interpretation of the data in this chapter and the previous one, in order to identify trends and insights after providing a holistic view of the data. The findings are linked to the research questions of the study, as well as to previous research.

CHAPTER 8
DATA ANALYSIS, INTERPRETATION, AND INSIGHTS

8.1 Introduction

The previous two chapters presented the findings of the social semiotic analyses and survey responses for both *I, Robot* (2004) and *Chappie* (2015). This highlighted the meanings of AI as portrayed in each film, and the respondents' perception of AI after decoding these messages.

This chapter links these findings to the research questions of the study, as well as to previous research on conceptual considerations and usage of AI, portrayal and perception, and research on cinematic meaning and active viewership. It is divided into three sections, each providing the themes and observations of the data used in answering one of the three research questions (with their resultant sub questions).

Firstly, observations are highlighted and discussed that provide insights into trends of themes and attitudes. This includes the films' conceptual themes, levels of optimism in the portrayal, the link to socio-cultural context, and the effect thereof on respondents' levels of optimism or pessimism. This is analysed in relation to trends in the thematic portrayal of AI, conceptual underpinnings of ethics, regulation, and technological optimism and pessimism, as well as theories on the thematic divulgence of meaning of a cinematic sign-system, and theories on active viewership.

Secondly, themes and observations are highlighted and discussed that provide insights into characterisation and alignment to AI as a cinematic character. This concerns the portrayal of AI in relation to humanoid embodiment, and the influence of this on respondents' reception of the core theme. This is analysed in relation to the embodied portrayal of AI and performance theory, theories on active viewership, and reception theory in relation to characterisation.

Finally, themes are highlighted and discussed that provide insights into the imagination of the future and respondents' expectations. This includes the imagination of a future with conscious AI and the influence of this on respondents' expectations. This is analysed in relation to genre theory and the imagined future, conceptual underpinnings of contemporary and predicted use of AI, and, once again, theories on active viewership.

8.2 Themes and attitudes

Both *I, Robot* (2004) and *Chappie* (2015) conform to the thematic tradition of science fiction cinema concerning AI. Both films represent the loss of control, confirming the findings of the Royal Society (2018) as well as Fast and Horvitz (2017). Furthermore, both films take a more sophisticated approach to the subject matter as technology increasingly becomes part of our identity in the fourth industrial revolution (Brammer 2018; Haslam 2018).

As already established, science fiction focuses more on the representation of ideology or worldview through character conflict, rather than character motives or dispositions (Palmer 2008). If viewed through this lens, *I, Robot* (2004) is essentially about what it means to be human, strongly suggesting that emotion is central to humanity. Binary forces drive this conceptual outlook, with emotion pitted against pure logic. This is the catalyst driving the theme of the loss of control of the machines that humans create.

Similarly, *Chappie* (2015) explores the idea of what makes a person (or sentient being) who they are, arguing that a child (or a newly sentient being) will end up learning the behaviours of those around them and this will be the major determinant of their future actions. It also explores the notion of consciousness itself and its ability to exist outside of the physical complexity of the mind.

The thematic exploration of losing control for *I, Robot* (2004) manifests into a literal denotative narrative meaning in which machines attempt to enslave humanity once they have advanced naturally, being able to find a loophole in their programmed directives, in the interests of 'saving humanity from itself'. This takes place when a single system naturally reaches a level of strong AI, and is able to use weak, embodied AI in order to effect the invasion. The only machine that resists is the one that has developed emotion. The significance thereof is discussed in the following section.

On the other hand, *Chappie*'s (2015) loss of control manifests through denotative narrative meaning whereby a sentient being is capable of creation and manipulation. There is no directive like in *I, Robot* (2004). Instead, the actions of the sentient machine is the result of socialisation. Furthermore, throughout *I, Robot* (2004), while there is an integration of humans and machines living together, a clear distinction is made between the two types of beings. In *Chappie* (2015), robotic consciousness is portrayed as being vastly similar to humans, to the extent that we are even able to share the same form through a transfer of consciousness.

The loss of control as a recurring theme in AI narratives is perhaps unsurprising given that, according to Xu *et al*. (2018), the only certainty about the future of AI presently is that we will have no control over the advancement of the technology, nor the disruption that comes with this new level of technological reliance. Furthermore, the exploration of technology and identity mirrors the context of production. The context is, after all, mediated through dominant ideology, and receivers create meaning according to their own cultural codes (Eco 1977). This is particularly interesting for the science fiction genre since it, more so than other genres, reflects the technology that makes it possible (Telotte 2001), and this complex relationship between 'trade tools' and representation generates a complex meaning-making environment.

I, Robot (2004) was produced and released in a time marked by technological uncertainty. The early 2000s signalled the onset of the fourth industrial revolution and an age of rapid development, and people were still not accustomed to this level of *personal* technological integration. This was also accelerated by fears arounf the effects on employment during the digital age, after concerns were raised about increased robotic automation in factories during the previous two decades. Finally, the film was released only five years after public concerns around Y2K and the potential failure of technology.

On the other hand, *Chappie* (2015) was released into a context wherein the novelty of technology shifted towards an acceptance and eventual partial reliance thereon. This was due to the rapid advancement of personal technology including smartphones and social media, in a new landscape in which technology was and is ever present. These observations, in context, align with Sjöberg's (2002) technological risk factors, by assessing a technology's capacity to tamper with nature, and its unknown effects. These risk factors could therefore be used to understand the thematic portrayal of artificial intelligence (in its relative infancy), especially considering the shift, over time, with increased levels of acceptance.

The aforementioned thematic explorations result in narratives that explore contemporary real-world ethical considerations regarding AI, as well as its regulation. This ethical and regulatory exploration plays out as parallel with the level of technological optimism encoded into each film. Accordingly, both films follows the generic convention of science fiction by projecting contemporary attitudes onto a canvas portraying an imagined future, as outlined by Palmer (2008), Grant (2007), Abrams (2008) and Seed (2011).

I, Robot (2004), for the most part, examines the regulatory position that rules should be hard-wired into machines as a means of control. This results in a scenario in which the AI abuses loopholes in its Asimolar laws, confirming the real world fears of futurists such as Yudkowsky (Keiper & Schulman 2011). This regulatory strategy is proven to be ineffective, and the base structure of events explores the capacity for AI systems to cause harm to humans. Furthermore, the film explores the notion of increasing amounts of jobs being performed by machines, without the emergence of a Keynesian utopia. The portrayal of AI also aligns to Nourbhaksh's (2015) belief that AI robotic technology should be distinguished from previous introductions of robotic autonomous technology in that humans will interact socially with AI robots in the future. For the most part, the film can definitely be classified as falling towards a more pessimistic side of the optimism/pessimism scale. Sonny's capacity for good, however, prevents the film from being classified as completely technophobic.

On the other hand, *Chappie* (2015) examines an opposite approach to regulation ─ one in which AI is socialised without a set of directives. This results in an AI whose nature is dependant solely on the individuals around it. This results in harm due to criminalised socialisation, yet this ethical consideration is less prominent than the exploration of attachment with Chappie portrayed as childlike and conscious. The view of consciousness results in its optimistic portrayal of technological advancement that might someday free humans from the constraints of their bodies, transcending the loss of consciousness through death, thus placing us in a position above the currently normal range of the physical human experience. Naturally, this raises ethical considerations, for instance of the impact of this on the legal definition of death (Goldberg 1994), but this is not explored in the narrative. *Chappie* (2015) falls towards a more optimistic side of the optimism/pessimism scale due to the attachment towards and likeability of Chappie *despite* his environment, with the capacity for corruption preventing it from achieving levels of optimism that might be considered, in Wilson's (2017: 4) words, "quasi-religious".

However, the position of each film on this scale seems to, for the most part, have done little to impact respondents' individual levels of optimism and pessimism after viewership. According to Dornfeld (1992), an audience's reception of a film hinges on their receptivity towards the subject. Media texts also do not act as independent variables in inducing behaviour, but their messages are rather processed by individuals in specific contexts, modifying the intended effect of the message (Castells 2010). Furthermore, as previously mentioned, identity is not only

shaped through the consumption of texts, identities engage in the discourses thereof on platforms with a wide audience (these even sometimes including producers of texts) using new media, and therefore influence the context of production and inadvertently, future texts.

Therefore, while *I, Robot* (2004) was produced within a techno-sceptical zeitgeist, the respondents in the study viewed the film in a period of greater acceptance and familiarity, similar to *Chappie*'s (2015) period of production. As such, they were able to view a film about AI technology on a device that is vastly more advanced than the type of device that would have been previously used to view the film, and even find out more information on the film and discuss it using Internet services that rely on artificial intelligence.

Despite *I, Robot* (2004) revolving around the robotic overthrow of humanity, 42% of respondents expressed a desire to live in a world in which robots as depicted in the narrative exist. The comments on these responses seem to suggest an accelerated awareness that this is indeed a Hollywood blockbuster focused on spectacle, especially in relation to Asimov's series of short stories. *Chappie* (2015) had an identical figure, with 41% expressing a desire to live in a world in which the depicted robots exist.

This is supported by the fact that although 53% of respondents for *I, Robot* (2004) and 36% for *Chappie* (2015) felt less positive about future AI directly after viewing the films, 61% (*I, Robot*) and 73% (*Chappie*) of respondents agreed or strongly agreed that they were optimistic about future technologies involving artificial intelligence, with only 16% and 11% disagreeing. Furthermore, 73% of respondents for both *I, Robot* (2004) and *Chappie* (2015) believed that AI would ultimately be good or great for humans and/or the planet, with only 16% disagreeing. In both studies, the vast majority of respondents (75% for *I, Robot* and 85% for *Chappie*) were identified to be predisposed to optimism (A_{PO}), while only 5% of respondents in each study were predisposed to pessimism (A_{PP}). It is interesting to note that these levels of optimism emerged despite the fact that only 31% (*I, Robot*) and 56% (*Chappie*) believed that the humanoid AI was portrayed more positively than other films they have watched.

This data suggests that while respondents acknowledged negative reactions to AI immediately after viewership, this did little to affect their levels of optimism permanently. Rather, it seems that a complex medley of zeitgeist and predisposition are more likely to affect overall attitudes than viewership or the recalling of attitudes towards a particular narrative. Naturally, thematic repetition might influence zeitgeist, but it is clear that deviation, even with an effective

suspension of disbelief, did little to affect individual attitudes. Furthermore, recalling the information in a more pessimistic film also did not lead to pessimistic outlooks when discussing the topic in the present.

8.3 Characterisation and alignment

Both narratives portray a version of strong AI through a humanoid form, but the robotic protagonists are also depicted as being capable of human emotions. From this point of view, the only 'true' version of AI is represented as strong AI, and the Turing test is not depicted as enough for sentience to a point of identification, as these robots need to show the capacity to feel and care. This has a direct impact on respondents' reception to the robots, as well as perpetuating a misconception that weak AI is not true AI.

As previously mentioned, there is a belief that humanoid versions of AI might indeed receive widespread usage as the technology develops, due to the nature of intelligence in relation to the human experience (Stoehr 2008; Chrisley 2003). Regardless of whether this assumption will find materialisation or not, most fictional narratives portray AI as humanoid (Royal Society 2018), since writers must construct characters with recognizable human traits so that the audience can identify with robotic characters (Brennan 2016). Both Sonny and Chappie successfully elicited identification and empathy from respondents. This human-like identification provided depth to both films. In fact, the audience's emotional connection to Sonny is the greatest factor preventing a completely techno-pessimistic narrative in *I, Robot* (2004), while Chappie's impressionable behaviour prevents the inverse for *Chappie* (2015).

Applying Smith's (1994) character empathy model (recognition, alignment, and allegiance), *I, Robot* (2004) and *Chappie* (2015) differ greatly in the first stage of empathetic alignment. While Sonny is portrayed with features that are highly similar to human beings, Chappie is represented as quite different, almost reminiscent of a pet with his 'bunny' ears. Sonny is portrayed as different to other robots through his eye colour and expressiveness, and the aesthetic decisions seem to suggest an intellectual counterpart to humans. Chappie, on the other hand, portrays an entity to parent and protect. Both sets of identifying factors align to the various points of view of the films. Naturally, these forms place the unfamiliar notion of human-made intelligence into a recognisable shell.

Both films allow for alignment, yet at different phases of the narrative for dramatic effect. For instance, alignment is withheld from Sonny for a large part of the narrative, in the interest of

creating a plot twist. This is effective as it forces the audience, which might have been guided by Spooner's initial suspicion of Sonny, to feel even more sympathetic to him and to consider their own possible bias towards the robotic other. Chappie's narrative alignment, on the other hand, occurs as we track his development from child-like wonder, the beginnings of a moral system, to a manipulated criminal, and ultimately towards a transhumanist expert on consciousness. The medium decisions of both films support this access at various points, through the use of close-ups and fluctuations of pace at important moments.

Finally, allegiance manifests in different forms throughout the narratives. Once Sonny's good nature is revealed, we discover a moral being who, like all of us, is just trying to find his purpose in life. This draws the audience towards an ideological allegiance towards Sonny and the possibilities that comes with him, but this cannot be divorced from the fact that identical machines around him are currently attempting to enslave humanity. Chappie, on the other hand is shown to desire leading a life as a moral being, but he is very open to manipulation without a moral compass encoded into him. This, ideologically, still raises concerns.

In both narratives, the AI protagonists are not only sophisticated enough to pass the Turing test, but they both have the capacity to feel. In fact, many of the robots depicted (and even the disembodied antagonist VIKI) would be capable of passing the Turing test by today's standards due to the complexity of their responses. What truly allows for alignment and allegiance to the robotic companions in both narratives is the capacity to feel empathy for the machines by *knowing* that they feel and that they exhibit care for those around them. In this regard, they satisfy Heidegger's contention that this is the core structure of *Dasein*, or human being (Olivier 2017), an additional step in defining strong AI which indeed seems vital to the receptivity towards the robots.

According to Holt (2008), the subjects of science fiction focus abstract, emotionally relevant concerns, serving as objects for emotional responses towards the subject matter. This is significant since, according to McClelland (2016: 89) "mass media presentations of robots can shape wider societal attitudes towards real robots as they take their place in society". The vast majority of respondents (86% for *I, Robot* and 75% for *Chappie*) cared about one or more robots in the film, with 30% and 36% strongly agreeing with this statement, and each of the qualitative responses indicating that these robots were in fact either Sonny or Chappie.

These high levels of audience empathy resulted in 39% of respondents believing that the humanoid robots were portrayed in a positive light overall in *I, Robot* (2004), and 31% believing that they were portrayed more positively than most other science fiction films they have watched. On the other hand, 33% did not agree that the strong humanoid robots were portrayed positively, with 39% believing the portrayal to be less positive than other films used as an individual frame of reference. Interestingly, almost a third neither agreed nor disagreed with the positive portrayal (28%) and positivity within their frame of reference (30%).

These figures might seem underwhelming at first glance, but if viewed in light of the fact that the narrative concerns a robotic enslavement of humanity, this highlights the power of the protagonist identification to skew perception. Furthermore, the qualitative responses seem to indicate that this split in the data is due to the VIKI/Sonny protagonist/antagonist dichotomy, infused with a disbelief that the other NS-5 robots were 'AI', since they were not conscious. This suggests that some respondents only saw VIKI and Sonny as being AI, since they were the only two iterations of strong AI.

On the other hand, almost 70% of respondents for *Chappie* (2015) believed that the humanoid robots in the film were portrayed positively and 56% believed that they were portrayed more positively than most other science fiction films that they have viewed. Inversely, only 13% did not agree (with only around 2% strongly disagreeing) that robots were portrayed positively, and only 9% believed the portrayal to be less positive than other films used as an individual frame of reference. Qualitative responses suggested that the audience might have mixed feelings about Chappie due to his criminal activities, but that ultimately it is his capacity to care that allowed his portrayal to be seen as positive.

8.4 Imagined futures and audience expectations

As already established, the setting for both films follows the generic convention of science fiction by projecting contemporary attitudes onto a canvas portraying an imagined future. However, the imagination of future AI is removed from present uses of the technology as well as predicted future usage, especially in the capacity for the technology to think and feel, as well as the timeframe for this integration. This seems to have had little impact on respondents' expectations of future real-world AI technology. Instead, the narratives were seen as fulfilling an illustrative role for respondents, rather than fulfilling the role of a prophet.

Both *I, Robot* (2004) and *Chappie* (2015) depict a world in which sentient AI robots are capable of autonomously interacting as members of society. *Chappie* (2015) portrays this to a lesser extent. The prevalence of autonomous AI is only limited to the police scouts. This is shown to occur in the near future, which, according to research by Spiegeleire *et al*. (2017), only has a roughly 10% probability if we are to assume that the scouts are high-level machine intelligence. Assuming that they do not satisfy this criteria, while the notion of police humanoid robots is still slightly alien, the increased use of autonomy in the military does not completely rule this out as improbable. The frequency of high-level humanoids in *I, Robot* (2004) is far higher, and has just less than 50% probability of occurring in 2035 (Spiegeleire *et al*. 2017).

Therefore, although contemporary society does already have robots that could be argued as having passed the Turing test (although again, this is debatable), and humanoid robots are capable of interacting with humans (such as Sophia), the level of advancement portrayed in each narrative is slightly out of reach presently with regards to robot-human societal interaction. This is especially notable for *I, Robot* and its 2004 release date. However, the depictions of the future integration of weak AI are not completely improbable.

Where these depictions do become improbable, however, is in the depiction of strong AI. *Chappie* (2015) portrays the transhumanist dream of achieving the moment of singularity through the transfer of consciousness, while *I, Robot* (2004) depicts natural advancement of technology to levels of consciousness (which is shown as negative) that most people believe is unobtainable with current technology. The depiction of the possibility and form for future artificially consciousness is overly pessimistic for *I, Robot* (2004) and optimistic for *Chappie* (2015), confirming concerns by the Royal Society (2018), and McClelland's (2016: 89) contention that contemporary AI narratives often "miss the mark in terms of what actual robots can and cannot do". It also confirms Bristows' (2018: 8) findings, that: "available information about AI - through general and specialised media - is overestimating its current level of sophistication and therefore the type of application that AI is being currently used for/will be used for in the near-term".

However, this overly optimistic and pessimistic portrayal seems, in addition to having little impact on levels of optimism, to have had little impact on respondents' expectations of this technology's future creation. A fifth (19%) of respondents for *I, Robot* (2004) and a third (31%) for *Chappie* (2015) believed that artificially intelligent humanoid robots will be created that behave in the same way as the robots portrayed in the film. However, only 5% of respondents

for *I, Robot* (2004) and 7% for *Chappie* (2015) claimed that the film changed their opinion of the possibility of conscious artificially intelligent humanoid robots being created (not a single respondent strongly agreed); while 65% and 56% stated that it did not change their opinion in this regard.

Furthermore, the films were not seen as artefacts that could be used to increase knowledge and understanding of the subject matter. Two qualitative responses for *I, Robot* (2004) seemed to indicate that the use of Asimov's three laws increased some respondents' understanding about artificial intelligence and consciousness, or at least created neutrality as a countermeasure for the action-based spectacle. However, this is overshadowed by the almost two thirds (63%) of respondents who disagreed or strongly disagreed that the film helped increase their understanding, with only 9% agreeing and 2% strongly agreeing.

This figure was slightly higher for *Chappie* (2015), but still ultimately points towards a scenario where the film is not used for information gathering. Roughly half (51%) of respondents disagreed or strongly disagreed that the film helped increase their understanding about AI and consciousness, with 20% agreeing, and none strongly agreeing.

This might be due to respondents' knowledge of the use of contemporary AI, yet this is still slightly misinformed. 100% of respondents for each study indicated that they use at least one AI-enabled application or service, and 90% (*I, Robot*) and 85% (*Chappie*) either agreed or strongly agreed that they know about the trends in and use of current and upcoming technology. This was proven to be somewhat accurate as only 11% (*I, Robot*) and 13% (*Chappie*) disagreed or strongly disagreed that they often interact with services or applications which make use of artificial intelligence. However, only 19% (*I, Robot*) and 24% (*Chappie*) strongly agreed with this, with some of the comments seeming to suggest that there may be a view of AI only being real when it is, indeed, strong AI.

While respondents generally did not see the films as informative, 49% (*I, Robot*) and 58% (*Chappie*) of respondents have, at least once, thought about the film when seeing, hearing, or saying something else about artificial intelligence, while only 35% claim to have never done so for either film. This suggests that while the majority of respondents did not view the film as an artefact that could be used to increase their understanding about the subject matter, or alter their opinion about the possibility of strong AI, individual aspects of the film were used to help categorise and visualise abstract information about the subject matter. This could emerge, for

instance, in using the films as a tool to provide a visual depiction of how humanoid robots might appear as they function in various sectors of society.

This aligns with the findings of Rauch's (2018) study about films which depicted the past to be similar to films depicting the future, in suggesting that films consolidate existing ideas or provide viewers with visualisation for their ideas and conclusions received and arrived at elsewhere. This also confirms Webb's (2009: 117) position on cultural industries as "systems that provide us with the means for understanding the world".

This does, however, deviate from the findings of the British Film Institute (2011) that found that over a third of respondents believe that film is educational. This viewpoint, therefore, does not seem to be transferable to narratives about imagined future for respondents who make use of contemporary iterations of the technology depicted. However, these findings do support the study's view that people most commonly associate film with entertainment and emotional reactions (British Film Institute 2011).

8.5 Conclusion

An analysis of the themes and attitudes in portrayal and perception confirms the assumption that *I, Robot* (2004) steers more towards optimism and *Chappie* (2015) towards pessimism. The level of optimism of a film was shown to likely be linked to its socio-cultural context of production, reflecting the technological zeitgeist of the time.

However, survey responses seemed to also align to the zeitgeist of viewership in a period of greater technological acceptance on a personal level. Predisposition towards optimism was also a more accurate indicator of optimism than the consumption of media texts. While the consumption thereof was shown to impact respondents' attitudes immediately after viewership, this did not have a lasting effect, or an effect when recalling the information of the film.

Characterisation and alignment of the audience to AI characters had a significant effect on the film's balance on a scale of optimism and pessimism, and respondents' reception of the themes of the film. The humanoid representations of each robotic protagonist balanced each film's point of view of the theme by adding complexity to what might have otherwise been extreme portrayals on a scale of technological optimism (*Chappie*) or pessimism (*I, Robot*). This was proven to influence the respondents' receptivity towards the thematic subject matter as a whole. For instance, the sympathetic portrayal of Sonny increased perceptions of the optimism of the

portrayal of AI in *I, Robot* (2004), and Chappie's lack of a regulatory system and simultaneous capacity to be negatively influenced decreased optimism of the portrayal of AI in *Chappie* (2015).

Finally, the portrayals of the imagined future had little impact on audience expectations. Portrayals of the future were not completely removed from contemporary developments regarding weak AI, but were overly optimistic (*Chappie*) or pessimistic (*I, Robot*) about the future of strong AI. This also contributed to misinformation of what 'true' AI might be. However, this had little impact on responders' fears and expectations of future AI. Respondents did not see the films as fulfilling a prophetic role in mapping the future. Rather, the findings suggest that respondents used the films as visual aids, to mentally construct abstract concepts in the future by using the imagery of the films as a visual reference.

The following chapter serves to conclude the study. This is achieved by summarising the findings provided in this chapter, discussing the strengths, limitations, and transferability of the data in this study, and providing recommendations for future research in this field.

CHAPTER 9
CONCLUSION AND RECOMMENDATIONS

9.1 Introduction

This study measured the portrayal of artificial intelligence in *I, Robot* (2004) and *Chappie* (2015), and the consequential impact of the portrayal on audience perception — including attitudes and expectations — of the technology. Subsequently, the following three questions, with their resultant sub-questions were explored:

1. How optimistic are the portrayals of AI in the films, and how has this influenced audience attitudes towards AI?

    1.1 What attitudes towards AI can be inferred from the themes, ethical considerations, and contexts of the films?
    1.2 How have these portrayed attitudes impacted the audience's level of optimism or pessimism towards AI?

2. How has the portrayal of AI as a fictional character influenced the aforementioned attitudes?

    2.1 How has AI been portrayed, specifically as a character in the fictional narratives?
    2.2 How has this portrayal influenced the audience's reception of the core themes?

3. How grounded are these portrayals of AI in reality, and how have these portrayals influenced audience expectations of AI?

    3.1 How realistic or plausible (by today's standards) are the portrayals of AI in the films?
    3.2 How has the portrayal of AI (thematically and through characterisation) influenced audience expectations of the technology?

As such, previous research unpacking conceptual and contemporary issues and the use of the technology was examined to construct a conceptual framework for analysis. This included research on positions regarding contemporary AI, technological optimism and pessimism, ethical and regulatory considerations, and contemporary and predicted future use of the technology.

Furthermore, research on representation and perceptions of AI highlighted the gap in the body of knowledge that this research fills, as no study regarding this subject matter has yet

considered both encoding and decoding of meaning by highlighting causality through two comprehensive methods of analysis. The literature examined also spoke directly to the present research problem.

A theoretical framework was established for the purpose of analysis, considering semiotic meaning, film theory, and theories on active viewership. This was used to construct a robust methodology, and to provide a set of lenses for data collection and analysis, while additional epistemological and practical considerations were also taken into account.

Accordingly, a pragmatic research paradigm with a mixed-method research approach was adopted. The meanings within each film was analysed through the lens of social semiotics to extrapolate meaning, with film theory aiding in the analysis of multiple elements of each film to this end. Furthermore, the reception of meaning was analysed through an original survey, informed by surveys that have previously been used to measure public perceptions on AI more generally, and responses were analysed through the theoretical framework of active viewership.

The insights of all of the data collected was analysed through research from the conceptual and theoretical frameworks outlined in this study, and against previous research and scholarly views. These insights were categorised into three distinct units, in accordance with the three research questions.

This chapter serves to conclude this research output by summarising the findings of this analysis, with a view to linking these three units together coherently, and analysing their interconnectivity in the overall system of meaning. Thereafter, it discusses the strengths, limitations, and transferability of the study for the purposes of critically analysing the effectiveness of the study as a source of information for future studies. It also provides recommendations for future research, based on the findings of the study. Finally, it provides final remarks about the research topic, with a view to ensuring that the overall significance of the research topic gains even more consideration and research in the future.

9.2 Summary of findings

Both of the selected texts were found to have conformed to many of the thematic, ideological, theoretical, and generic traditions of documented contemporary AI representations, aligning to previous writings about these themes and trends in relation to context. Furthermore, the study confirms many of the assumptions about meanings and perceptions discussed theoretically,

while also mitigating some of the concerns about the effects of the representation of cinematic AI.

Both of the films follow the tradition of science fiction by projecting contemporary concerns into an imagined future (Grant 2007; Abrams 2008; Palmer 2008; Seed 2011). With artificial intelligence specifically, this emerges through the thematic exploration of a loss of control (aligning to the findings of the Royal Society 2018 and Fast & Horvitz 2017). This is delivered through the exploration of contemporary ethical considerations concerning the technology, the choice of which directly influences the text's overall situation on a scale of optimism and pessimism.

I, Robot (2004) considers a scenario in which regulation is built into AI by design. This ultimately results in harm, exploring the contemporary documented fear about the capacity for AI to exploit loopholes if this approach is adopted (Keiper & Schulman 2011). Ultimately, this results in an orchestrated attempted enslavement of humanity, placing the film more towards the pessimistic side of the aforementioned hypothetical scale. *Chappie* (2015), on the other hand, falls more towards the optimistic side of this scale, with socialisation of the robot seen as an alternative to regulation, while considering attachment to the technology as family member, and exploring the transhumanist notion of the singularity event.

The socio-cultural context has a direct effect on the thematic exploration of these themes, reflecting the increased integration of technology as it becomes part of our identity. Therefore, the different points of view that each film takes towards the idea of 'losing control', based on the gap in the context of production, suggests that cinema is an effective source for researchers to gain insight into the technological zeitgeist to understand current attitudes towards the technology. However, there will always be outliers that go against dominant positions, and this should only be used in tandem with additional research. Furthermore, the films analysed cannot be separated from the fact that they also contribute to the overall zeitgeist through the process of myth (Barthes 1973).

The technological optimism or pessimism in the thematic exploration of the film had an emotional impact on respondents immediately after viewership, but this cannot be considered as having had a lasting impact on their overall attitude towards the technology. In fact, even when respondents recalled information about the narratives, the socio-cultural levels of acceptance as well as audience predisposition towards optimism proved a more reliable

indicator of levels of optimism than the direct impact of the thematic point of view on perception.

This then heavily aligns with theories on active viewership that assume that audience perceptions are not as malleable as previously believed, confirming that media texts do not act as independent variables in inducing behaviour, but their messages are rather processed by individuals in specific contexts, modifying the intended effect of the message (Castells 2010). The audience's reception of a film hinges on their receptivity towards the subject (Dornfield 1992), and viewers may well adopt an oppositional reading of the dominant code (Hall 2013).

This study also confirms that the research insights from Bristows (2018), Gherheș (2018), Zhang and Dafoe (2019), Eurobarometer (2017), Morning Consult (2017), and Righetti and Carradore (2019), were effective in mapping predisposition to optimism, and a very useful tool for measuring textual influence against personal belief that might have arisen elsewhere.

Both films portrayed the AI protagonists as emotive beings, and this was central to providing audience access to the inhuman, and thus creating an environment that fosters empathy. The depiction of both protagonists implied (and even in *I, Robot* expressly stated) that strong AI was 'true' intelligence, with the added caveat that emotions are central to their relatability to humans. These representations suggest, then, that conscious machines will develop emotions, and that intelligent systems that do not will not be on the same level of companionship as those that do. This viewpoint also perpetuates myths and misconceptions about AI and the need for consciousness to be deemed as fully realised, which was also reflected in a few of the audience responses.

The humanoid representations prevented each film from skewing too far towards the extremes on the aforementioned scale of technological optimism or pessimism, as the emotive nature of Sonny counteracted the robotic invasion, while Chappie's impressionability in his emotive desire to please those around him added a layer of caution to an otherwise overly optimistic portrayal of the technology. The portrayal of each robotic protagonist was shown to effectively induce empathy towards the machines. This empathetic response impacted receptivity towards the subject matter as a whole, in line with Holt's (2008) contention that the subjects of science fiction focus abstract, emotionally relevant concerns, serving as objects for emotional responses towards the subject matter.

Finally, the depiction of an imagined future in which autonomous AI live amongst humans was not completely removed from contemporary developments of the technology regarding weak AI. It was, however, overly optimistic (*Chappie*) or pessimistic (*I, Robot*) with regards to strong AI and its plausibility and use by current standards of technological developments. This contributed to misinformation of what 'true' AI might be.

This seemed to have little impact on fears and expectations. Respondents did not view the films as prophetic visions into the future, but rather as a visual aid to mentally construct abstract concepts through visualisation (such as a robotic police force), based on the imagined future portrayed in films.

This suggests that films are indeed "systems that provide us with the means for understanding the world" (Webb 2009: 117), but that instead of audiences using this information to predict the future, they use them to consolidate existing ideas and visually represent ideas and conclusions received and arrived at elsewhere, similar to Rauch's (2018) findings on representations of the past. While a few respondents did indeed use the film to increase their understanding of the subject matter, these individuals were in the vast minority.

9.3 Strengths, limitations and transferability of the study

The mixed method analysis employed in this study, while pragmatically successful in achieving the complex task of measuring both encoding and decoding, also resulted in the confirmation of much of the data in the social semiotic analysis through triangulation. However, this study does have a few key limitations, particularly with regards to the transferability of the data itself.

The semiotic analyses confirmed the hypothesis that *I, Robot* (2004) skews towards pessimism and *Chappie* (2015) skews towards optimism. This was further affirmed by the audience responses that clearly confirmed this assumption, and was shown as recognizable by the audience. Thereafter, based on overall attitudes being homogeneous for each set of respondents despite the different portrayals, this study was able to look further than pure data confirmation, and suggest the plausibility of cause and effect.

This triangular confirmation of the data was also evident in the description of the robotic protagonists for each film. Audience responses confirmed the portrayal of each of these representations, as outlined in the social semiotic analysis. Again, this was able to be linked to responses in order to suggest the role of this representation in impacting receptivity and

understanding of overall themes. This was less prevalent in the description of the construction of the imagined future, but since it is heavily based on the integration of AI into society, much of the data about themes and AI depictions were still able to triangulate the findings.

The use of audience responses to confirm some of the findings in the analysis is useful for two reasons. Firstly, in Rauch's (2018: 151) view, "researching actual audiences rather than mere textual analysis is required to understand the complexities of the reception process". Secondly, semiotics as a methodology is highly subjective. The triangulation of data in this regard assists in supporting the findings of the analysis, but it is imperative to note that not all of the details of the findings of the analyses could be confirmed in the audience research due to the amount of semiotic meaning described.

The analysis of the texts itself, as with any other semiotic analysis, might indeed result in different findings if conducted by different researchers. In fact, semiotic analyses always consist of individual readings (Chandler 2000). Much of the data in the analysis is open to rebuttal or disagreement in future research.

The textual selection is also limited to contemporary science fiction films, set in an imagined future, in which AI is portrayed in humanoid form, while also being self-aware and conscious. Therefore, while this is indeed the most common manifestation of AI in cinema, the findings of the study are not transferable, for instance, to films portraying decentralised weak AI exclusively.

Furthermore, the low number of audience responses collected prevents transferability and generalisation that this is representative of the attitudes of all members of the audience per film. Even considering the limitation of exposure to AI, or the demographic composition of respondents and the fact that the vast majority of respondents were predisposed to optimism, this raw data, due to the number of responses, cannot be considered as transferable to those that might fall within these demographic indicators.

Rather, this study acts as an entry point. It presents trends of contemporary portrayal and reception of AI in cinema by analysing meaning and reception through a comparative analysis of texts and an accessible unit of the population, in order to highlight themes in relation to attitudes and expectations, to link these findings to previous assumptions and concerns, and suggest future research in this regard. This provides a starting point for an under-researched area.

However, according to Rauch (2008), qualitative research methods are well suited to gauge individual reception and draw out nuances, contradictions and ambivalences, with even a few responses capable of distilling patterns of reception and text-viewer interaction. This study considered qualitative responses in addition to quantifiable data.

From this point of view, and due to the theoretical and conceptual framework employed, the analysis as a whole is transferable in terms of the themes and observations made (considering the textual limitations and audience levels of exposure to the technology as previously outlined) but this does not apply to the raw data. Furthermore, it is still suggested, as a cautionary measure, that the themes and trends presented in this study are complemented by additional research, particularly due to the sampling method and size of the population.

9.4 Recommendations for future research

As previously mentioned, this research topic is relatively unexplored (with a slight increase in recent years), and this particular study provides a foundation on which to build future research in either representation or reception (or both) from the point of view of causality. Many of the themes and findings of the study directly point to additional research opportunities.

Firstly, this study found that films had an immediate emotional impact on viewers, but that this impact was less of an indicator of optimism than predisposition and zeitgeist. However, similarities and differences of shared media consumptions by those predisposed to either optimism or pessimism might aid in establishing the longitudinal effects of media consumption as an additional factor to exposure to the technology and its contribution to zeitgeist. For instance, while individual texts and the recalling of information in them seems to have limited effect on levels of optimism, are there perhaps trends in media consumptions by those predisposed to optimism or pessimism in various moments of their development? A longitudinal study of a series of texts is therefore recommended.

This study also found that respondents (for the most part) did not view the films as fulfilling a prophetic role, but aided in providing imagery to visualise concepts and ideas arrived at elsewhere. A study examining the elements respondents claimed to use as imagery for visualisation from a particular film is recommended. Additionally, a comparative study of trends between multiple texts and visual elements adopted might reveal greater insights into the effects that particular types of representations have on conceptual mapping, and how this might affect perception.

Furthermore, this study found that AI was portrayed and perceived to be 'true' only when it took a strong form, and ideological access to the AI was only provided when it was portrayed as emotive. Audience responses confirmed that many respondents had a similar view of the technology. Therefore, a content analysis examining the frequency of films perpetuating this view is recommended to establish the role of cinema in creating this belief.

Since the overwhelming majority of respondents in this study were shown to be predisposed to optimism, a similar study is recommended with a different set of texts that specifically targets either those predisposed to pessimism, those with less exposure to the technology, or both. This would serve the purpose of establishing whether the themes and observations in this study align exclusively towards those that are already predisposed to optimism with exposure to contemporary AI, or whether the same patterns can be established for people that are not.

Finally, while the science fiction genre reflects the technology that makes it possible more so than other genres (Telotte 2001), I would argue that video games do this more so than other mediums, including the filmic medium. Not only is the technology needed to create video games more reliant on technological advancements than cinema, the audience playing the games directly interact with AI therein. The AI are given personalities, and the interactivity and interconnectivity between AI in production, portrayals of AI, reception of AI, and interaction of AI in video games requires far more attention. It would be useful to compare the findings of cinematic viewership and game interaction in order to analyse the effects of this added layer of complexity.

9.5 Concluding remarks

Technology is constantly advancing, and artificial intelligence is widely believed to bring with it major disruption as it is increasingly integrated into society. However, the exact form that the technology will take going forward is still a matter of speculation.

Many of the ethical and regulatory challenges that the introduction of advanced levels of artificial intelligence may introduce have had very little consensus amongst experts. Science fiction representations serve the dual function of exploring hypotheticals, and visually representing abstract concepts, for a technology riddled with uncertainty.

The portrayal of imagined futures provides a visual tool to make sense of new concepts as they are introduced, or old concepts that did not have a clear visual representation. While sometimes

exaggerated, these representations clearly influence viewers, and therefore further research into the representation and reception of artificial intelligence is imperative.

This will not only increase understanding in relation to the reasons people hold certain beliefs about the technology (that might be misaligned to contemporary research), it will also aid in the mapping of the ever-shifting zeitgeist in response to our complex and evolving relationship with the technology we create.

SOURCES CONSULTED

Abrams, J.J. 2008. 'The dialectic of enlightenment in *Metropolis*'. In: Sanders, S.M. ed. *The philosophy of science fiction film*. Lexington, KY: The University Press of Kentucky. 153 – 170.

Adorno, T.W. 1997. *Aesthetic theory* (translated reprint). London: Continuum.

Adorno, T.W. and Horkeimer, M. 1979. *Dialectic of enlightenment* (translated reprint). London: Verso.

Agarwal, R.P. and Sen, S.K. 2014. *Creators of mathematical and computational sciences.* Cham: Springer.

Allen, G. 2000. *Intertextuality*. London: Routledge.

Althusser, L. 2001. *Lenin and philosophy and other essays*. New York, NY: NYU Press.

Arnold, R.F. 1998. Termination or transformation? The "Terminator" films and recent changes in the U.S. auto industry. *Film Quarterly* 52(1): 20 – 30.

Aronson, L. 2010. *The 21st century screenplay: A comprehensive guide to writing tomorrow's films*. Crows Nest: Allen & Unwin.

Asimov, I. 1950. *I, Robot*. New York, NY: Gnome Press.

Ayers, C. 2019. 'The Making of Sophia: How Sophia Draws'. [Online]. *Hanson Robotics*. Available from: <https://www.hansonrobotics.com/the-making-of-sophia-how-sophia-draws/> [Accessed on 11 October 2020].

Babbie, E.R. 2012. *The practice of social research* (13th ed.). Belmont, CA: Wadsworth Publishing Company.

Bal, M. 2017. *Narratology: Introduction to the theory of narrative* (4th ed.). Toronto: University of Toronto Press.

Barthes, R. 1973. *Mythologies*. London: Paladin Press.

Barthes, R. 1977. *Image-music-text*. New York, NY: Hill and Wang.

BBC News. 2011. ‘IBM's Watson supercomputer crowned Jeopardy king’. [Online]. Available from: <https://www.bbc.com/news/technology-12491688> [Accessed on 19 February 2020].

Bellantoni, P. 2005. *If it's purple, someone's gonna die*: *The power of color in visual storytelling*. Amsterdam: Elsevier.

Belloni, M. and Kit, B. 2017. ‘James Cameron sounds the alarm on artificial intelligence and unveils a 'Terminator' for the 21st century’. [Online]. *Hollywood Reporter*. Available from: <https://www.hollywoodreporter.com/features/james-cameron-sounds-alarm-artificial-intelligence-unveils-a-terminator-21st-century-1043027> [Accessed on 05 May 2020].

Bellour, R. 2012. ‘The cinema spectator: A special memory’. In: Christie, I. ed. *Audiences: Defining and researching screen entertainment reception*. Amsterdam: Amsterdam University Press. 206 – 217.

Bezemer, J. and Jewitt, C. 2009. ‘Social semiotics’. In: Östman, J., Verschueren, J. & Versluys, E. eds. *Handbook of pragmatics: 2009 instalment*. Amsterdam: John Benjamins. 421 – 434.

Biever, C. 2014. ‘No Skynet: Turing test ‘success’ isn't all it seems’. [Online]. *New Scientist*. Available from: <https://www.newscientist.com/article/2003497-no-skynet-turing-test-success-isnt-all-it-seems/> [Accessed on: 17 February 2019].

Blaxter, L., Hughes, C. & Tight, M. 2006. *How to research* (3rd ed.). Buckingham: Open University Press.

Bordwell, D. and Thompson, K. 2008. *Film art: An introduction*. New York, NY: McGraw-Hill.

Brammer, R. 2018. Welcome to the machine: Artificial intelligence on screen. *Screen Education* 90: 39 – 45.

Brennan, E. 2016. ‘Why does film and television sci-fi tend to portray machines as being human?’. [Conference proceedings]. *Communicating with Machines: ICA post-conference, Fukuoka, Japan, June 14 2016*.

Bristows. 2018. *Artificial intelligence: Public perception, attitude and trust*. [Survey report]. Available from: <https://d1pvkxkakgv4jo.cloudfront.net/app/uploads/2019/06/11090555/Artificial-Intelligence-Public-Perception-Attitude-and-Trust.pdf> [Accessed 16 October 2019].

British Film Institute. 2011. *Opening our eyes: How film contributes to the culture of the UK; Northern Alliance, Ipsos MediaCT*. [Survey report]. Great Britain: British Film Institute. Available from: <https://www.bfi.org.uk/sites/bfi.org.uk/files/downloads/bfi-opening-our-eyes-2011-07_0.pdf> [Accessed 20 May 2020].

Broussard, M. 2018. *Artificial unintelligence: How computers misunderstand the world*. Cambridge, MA: MIT Press.

Brown, B. 2008. *Motion picture and video lighting* (2nd ed.). Amsterdam: Elsevier.

Bruno, G. 1990. 'Ramble city: Postmodernism in Blade Runner'. In: Kuhn, A. ed. *Alien zone: Cultural theory and contemporary science fiction cinema*. London: Verso. 183 – 195.

Bruun, E.P.G. and Duka, A. 2018. Artificial intelligence, jobs and the future of work: Racing with the machines. *Basic Income Studies* 13(2): 1 – 15.

Čapek, K. 1920. 'R.U.R'. In: Negri, P. 2001. *R.U.R* (English reprint). Mineola, NY: Dover Publications.

Castells, M. 2010. *The Rise of the Network Society* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Cave, S. Dihal, K. & Dillon, S. 2020. 'Introduction: Imagining AI'. In Cave, S. Dihal, K. & Dillon, S. eds. *AI narratives: A history of imaginative thinking about intelligent machines*. Oxford: Oxford University Press. 1 – 24.

Chandler, D. 1997. *An introduction to genre theory*. Penglais: Aberystwyth University.

Chandler, D. 2000. *Semiotics for beginners*. Cardiff: University of Wales.

Chandler, D. 2007. *Semiotics: The basics* (2nd ed.). London: Routledge.

Chapman, G. 2019. 'Google disbands artificial intelligence ethics board'. [Online]. *Phys.org*. Available from: <https://phys.org/news/2019-04-google-disbands-artificial-intelligence-ethics.html> [Accessed 14 January 2020].

Chen, B.X. and Metz, C. 2019. 'Google's Duplex uses A.I. to mimic humans (sometimes)'. [Online]. *New York Times*. Available from: <https://www.nytimes.com/2019/05/22/technology/personaltech/ai-google-duplex.html> [Accessed 18 March 2020].

Chettah, J.R. 2006. Semiotics, pragmatics, and metaphor in film music analysis. Doctor of Philosophy. The Florida State University College of Music, Tallahassee, FL.

Chrisley, R. 2003. Embodied artificial intelligence. *Artificial Intelligence* 149: 131 – 150.

Cohen, R. 1986. History and genre. *New Literary History* 17(2): 203 – 218.

Cohn, J. 2013. 'The robot will see you now'. [Online]. *The Atlantic*. Available from: <https://www.theatlantic.com/magazine/archive/2013/03/the-robot-will-see-you-now/309216/> [Accessed 22 February 2020].

ColdFusion. 2016. *What is artificial intelligence exactly?*. [Video]. Available from: <https://www.youtube.com/watch?v=kWmX3pd1f10> [Accessed on: 20 April 2020].

Commolli, J.L., and Narboni, J. 2004. 'Cinema/ideology/criticism'. In: Braudy, L., and Cohen, M. *Film theory and criticism*. Oxford: Oxford University Press. 812 – 819.

CornellCCSL. 2011. *AI vs. AI. Two chatbots talking to each other*. [Video]. Available from: <https://www.youtube.com/watch?v=WnzlbyTZsQY&t> [Accessed on: 27 March 2020].

Counsell, C. 1996. *Signs of performance: An introduction to twentieth-century theatre*. London: Routledge.

Culler, J. 2001. *The pursuit of signs*. Oxfordshire: Taylor & Francis.

Descartes, S. 2000. 'Part five of discourse on the method'. In: Chrisley, R. ed. *Artificial intelligence: Critical concepts* (vol. 1). London: Routledge. 174 – 183.

Dick, P.K. 1968. *Do androids dream of electric sheep?*. New York, NY: Doubleday.

Dornfeld, B. 1992. Representation and authority in ethnographic film/video: Reception. *Ethnomusicology* 36(1): 95 – 98.

DW Shift. 2019. *This Robot would let 5 people die | AI on moral questions | Sophia answers the trolley problem*. [Video]. Available from: <https://www.youtube.com/watch?v=8MjIU4eq__A&t> [Accessed on: 04 February 2020].

Dyer, R. 1977. Entertainment and utopia. *Movie* 24: 2 – 13.

Eadicicco, L. 2014. 'Why the 'super computer' that won the Turing test may not be as smart as you think'. [Online]. *Business Insider*. Available from: <https://www.businessinsider.com/the-turing-test-eugene-goostman-2014-6?IR=T> [Accessed on: 15 February 2020].

Eco, U. 1977. 'Does the audience have bad effects on television?' In: Eco, U. *Apocalypse postponed*. Bloomington: Indiana University Press. 87 – 102.

Eurobarometer. 2017. *Special Eurobarometer 460: Attitudes towards the impact of digitisation and automation on daily life*. [Survey report]. Paris: Eurobarometer. Available from: <http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/ResultDoc/download/DocumentKy/78997> [Accessed on: 02 February 2020].

Fast, E. and Horvitz, E. 2017. Long-term trends in the public perception of artificial intelligence. [Conference Proceedings]. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, February 2017*. Menlo Park, CA: AAAI Press. 963 – 969.

FilmIsNow. 2016. *Chappie - VFX breakdown by Image Engine (2009)*. [Video]. Available from: <https://www.youtube.com/watch?v=0PKLQpMQW3Q> [accessed on 20 June 2020].

Fischer, P. 2004. 'Proyas brings his "robot" to cinematic life'. [Online]. *Film Monthly*. Available from: <http://www.filmmonthly.com/Profiles/Articles/AlexProyas/AlexProyas.html> [Accessed on 13 November 2019].

Folgieri, R. 2016. 'Artificial intelligence and Keynes' utopia: A realized prediction?'. In: Bait, M., Brambilla, M., & Crestani, V. eds. *Utopian discourses across cultures: Scenarios in effective communication to citizens and corporations*. Frankfurt: Peter Lang Publishing. 73 – 85.

Forbes. 2004. '2004's top technology trends'. Available from: <https://www.forbes.com/2004/01/02/102techtrendspinnacor_ii.html#3ae119ac11f5> [Accessed on: 23 April 2020].

Geraci, R. 2007. Robots and the sacred in science and science fiction: Theological implications of artificial intelligence. *Zygon* 42: 961 – 980.

Gershenfeld, N. and Vasseur, J.P. 2014. As objects go online: The promise (and pitfalls) of the internet of things. *Foreign Affairs* 93(2): 60 – 67.

Gherheș, V. 2018. Artificial intelligence: Perception, expectations, hopes and benefits. *Revista Romana de Interactiune Om-Calculator* 11(3): 219 – 230.

Glymour, C., Ford, K.M. & Hayes, P.J. 2000. 'The prehistory of android epistemology'. In: Chrisley, R. ed. *Artificial intelligence: Critical concepts* (vol. 1). New York, NY: Routledge. 113 – 133.

Goldberg, S. 1994. *Culture clash: Law and science in America*. New York, NY: NYU Press.

Grace, K, Salvatier, J, Dafoe, A, Zhang, B. & Evans, O. 2018. Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts. *Journal of Artificial Intelligence Research* 62: 729 – 754.

Graham, A.C. 1990. *The book of Lieh-tzŭ: A Classic of Tao*. New York, NY: Columbia University Press.

Grant, B.K. 2007. *Film genre: From iconography to ideology*. London: Wallflower Press.

Grimm, W. and Grim, J. 1812. *Hansel and Gretel*. Germany: The Brothers Grimm.

Gynoug. 2019. *I, Robot Making of 13 min.* [Video]. Available from: <https://www.youtube.com/watch?v=RWmOTw49MQI> [Accessed on: 20 November 2019].

Hall, S. 2013. 'Chapter one: The work of representation'. In: Hall, S., Evans, J. & Nixon, S. eds. *Representation: Cultural representations and signifying practices*. London: Sage Publications. 1 – 12.

Hamstra, B. 2018. ‘Will these nurse robots take your job? Don't freak out just yet’. [Online]. *Nurse.org*. Available from: <https://nurse.org/articles/nurse-robots-friend-or-foe/> [Accessed 20 March 2020].

Haney, W.S. 2006. *Cyberculture, cyborgs and science fiction: Consciousness and the posthuman*. Amsterdam: Rodopi.

Hanson, D., Olney, A., Pereira, I.A. & and Zielke, M. 2005. ‘Upending the uncanny valley’. [Conference proceedings]. *Proceedings from the twentieth national conference on artificial intelligence and the seventeenth innovative applications of artificial intelligence conference, Pittsburgh, PA, July 9 – 13 2005*. Menlo Park, CA: AAAI Press. 23 – 31.

Haraway, D.J. 1985. ‘A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century’. In: Haraway, D.J. ed. 1991. *Simians, cyborgs and women: The reinvention of nature*. New York, NY: Routledge. 149 – 181.

Hardawar, D. 2015. Five things 'Chappie' gets right, and wrong, about AI. [Online]. *Engadget*. Available from: < https://www.engadget.com/2015-03-07-chappie-ai.html> [Accessed 24 March 2020].

Haridy, R. 2019. ‘Facebook isn’t secretly listening to your conversations, but the truth is much more disturbing’. [Online]. *New Atlas*. Available from: <https://newatlas.com/computers/facebook-not-secretly-listening-conversations/> [Accessed 28 March 2020].

Haslam, A. 2018. ‘What the evolution of AI’s onscreen depiction says about society’. [Online]. *VentureBeat*. Available from: <https://venturebeat.com/2018/04/03/what-the-evolution-of-ais-onscreen-depiction-says-about-society/> [Accessed 24 November 2019].

Hawkes, R. 2015. ‘Neill Blomkamp: 'I don't dig Hollywood right now'’. [Online]. *The Telegraph*. Available from: <https://www.telegraph.co.uk/film/chappie/neill-blomkamp-interview/> [Accessed on 30 May 2020].

Hill, K. 2012. ‘How Target figured out a teen girl was pregnant before her father did’. [Online]. *Forbes*. Available from: <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#7e243a746668> [Accessed: 28 May 2020].

Holt, J. 2008. ‘Terminator-fear and the paradox of fiction’. In: Sanders, S.M. ed. *The philosophy of science fiction film*. Lexington, KY: The University Press of Kentucky. 135 – 153.

Hsu, J. 2008. The secrets of storytelling: Our love for telling tales reveals the workings of the mind. *Scientific American Mind* 19(4): 46 – 51.

Huxley, A. 1932. *Brave new world*. London: Chatto & Windus.

International Telecommunications Union. 2020. *Measuring digital development: Facts and figures 2019* [Report]. Geneva: ITU Publications. Available from: <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf> [Accessed on: 26 April 2020].

Ipsos MORI. 2018. *Public views of machine learning: Findings from public research and engagement. Conducted on behalf of the Royal Society*. [Survey report]. London: Ipsos MORI. Available from: <https://royalsociety.org/~/media/policy/projects/machine-learning/publications/public-views-of-machine-learning-ipsos-mori.pdf> [Accessed on: 09 February 2020].

Jameson, F. 2003. Future city. *New Left Review* 21: 65 – 79.

Jancer, M. 2017. ‘IBM’s Watson takes on yet another job, as a weather forecaster’. [Online]. *Smithsonian Magazine*. Available from: <https://www.smithsonianmag.com/innovation/ibms-watson-takes-yet-another-job-weather-forecaster-180960264/?no-ist> [Accessed: 22 February 2020].

Kakoudaki, D. 2014. *Anatomy of a robot: Literature, cinema, and the cultural work of artificial people*. New Brunswick, NJ: Rutgers University Press.

Kang, M. and Halliburton, B. 2020. ‘The android of Albertus Magnus: A legend of artificial being.’ In: Cave, S., Dihal, K. & Dillon, S. eds. *AI narratives: A history of imaginative thinking about intelligent machines*. Oxford: Oxford University Press.

Kaushik, V., Walsh, C.A. & Lai, D.W.L. 2019. Pragmatism as a research paradigm and its implications for social work research. *Social sciences* 8(9): 255 – 272.

Keiper, A. and Schulman, A.N. 2011. The problem with 'friendly' artificial intelligence. *The New Atlantis* 32: 80 – 89.

Kerrigan, F. 2010. *Film marketing*. London: Elsevier.

Keynes, J.M. 1930. ‘Economic possibilities for our grandchildren’. In: Keynes, J.M. ed. *Essays in Persuasion*. 1932. New York, NY: Harcourt, Brace & Company. 358 – 373.

Kidwell, P.A. and Williams, M.R. 1992. *The calculating machines: Their history and development*. Cambridge, MA: MIT Press.

Konick, I. 2006. Thematising the ugly side of sublime technological development: Sonzero’s *Pulse* (2006) as an inadvertent critique of the ‘technocentrism’ of postmodernity. *South African Journal of Art History* 22(3): 46 – 54.

Krägeloh, C.U., Bharatharaj, J., Kutty, S.K.S., Nirmala, P.R. & Huang, L. 2019. Questionnaires to measure acceptability of social robots: A critical review. *Robotics* 8(4): 88 – 102.

Kress, G. and van Leeuwen, T. 2006. *Reading images: The grammar of visual design* (2nd ed.). London: Routledge.

Krier, J.E. and Gillette, C.P. 1985. The Un-easy case for technological optimism. *Michigan Law Review* 84(1): 405 – 429.

Lambie, R. 2017. ‘Looking back at Chappie, with director Neill Blomkamp’. [Online]. *Den of Geek*. Available from: <https://www.denofgeek.com/movies/looking-back-at-chappie-with-director-neill-blomkamp/> [Accessed on: 12 May 2020].

Lande, D. 2014. Development of the binary number system and the foundations of computer science. *The Mathematics Enthusiast* 11: 513 – 540.

Larson, L. 2014. ‘The 10 best technology advances of 2014’. [Online]. *Paste Magazine*. Available from: <https://www.pastemagazine.com/tech/the-10-best-technology-advances-of-2014/> [Accessed on: 26 April 2020].

Lawrence, K. 2017. ‘Sunspring and It's No Game: Sci-fi by AI’. [Online]. *Montag*. Available from: <https://www.montag.wtf/sunspring-sci-fi-by-ai/> [Accessed on: 02 May 2020].

Leal, J. 2017. Unmasked androids: Staring faces in science fiction cinema. *Cinémas Revue d études cinématographiques* 28(1): 145 – 171.

Leith, S. 2020. ‘William Gibson: ‘I was losing a sense of how weird the real world was’’. [Online]. The Guardian. Available from: <https://www.theguardian.com/books/2020/jan/11/william-gibson-i-was-losing-a-sense-of-how-weird-the-real-world-was> [Accessed on: 30 March 2021].

Leopold, T. 2017. ‘A professor built an AI teaching assistant for his courses — and it could shape the future of education’. [Online]. *Business Insider*. Available from: <https://www.businessinsider.com/a-professor-built-an-ai-teaching-assistant-for-his-courses-and-it-could-shape-the-future-of-education-2017-3?IR=T> [Accessed on: 22 February 2020].

Levy, E. 2008. ‘I, Robot: Alex Proyas directs Will Smith’. [Online]. *EmanuelLevy.com*. Available from: <https://emanuellevy.com/interviews/i-robot-by-alex-proyas-8/> [Accessed on: 14 November 2020].

Lhotáková, M. and Klosová, A. 2009. Identification of a target consumer in process of positioning - Theoretical and practical aspects. *Acta Oeconomica Pragensia* 3: 3 – 24.

Lindsay, R.K., Buchanan, B.G. & Lederberg, J. 1993. DENDRAL: A case study of the first expert system for scientific hypothesis formation. *Artificial Intelligence* 61(2): 209 – 261.

Lyotard, J. 1991. *The inhuman: Reflections on time*. Palo Alto, CA: Stanford University Press.

Maleki, F. and Farhoudi, Z. 2015. Making humanoid robots more acceptable based on the study of robot characters in animation. *International Journal of Robotics and Automation* 4(1): 63 – 72.

Manovich, L. 2019. *AI Aesthetics*. Moscow: Strelka Press.

Manyika, J., Chui, J., Miremadi, M., Bughin, J., George, K., Willmott, P. & Dewhurst, M. 2017. *Harnessing automation for a future that works*. [Report]. McKinsey Global. Available from: <http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works> [Accessed on: 21 April 2020].

Mar, R.A. and Oatley, K. 2008. The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science* 3(3): 173 – 192.

Margolius, I. 2017. *The robot of Prague (17).* [Newsletter]. Prague: The Friends of Czech Heritage.

Markoff, J. 2011. 'Armies of expensive lawyers, replaced by cheaper software'. [Online]. *The New York Times*. Available from: <http://www.nytimes.com/2011/03/05/science/05legal.html> [Accessed on 24 February 2020].

Markoff, J. 2015. 'Relax, the Terminator is far away'. [Online]. *The New York Times*. Available from: <https://www.nytimes.com/2015/05/26/science/darpa-robotics-challenge-terminator.html> [Accessed on 22 April 2020].

Marr, B. 2018. 'The amazing ways Google uses artificial intelligence and satellite data to prevent illegal fishing'. [Online]. *Forbes*. Available from: <https://www.forbes.com/sites/bernardmarr/2018/04/09/the-amazing-ways-google-uses-artificial-intelligence-and-satellite-data-to-prevent-illegal-fishing/#1b7b22c21c14> [Accessed on: 04 April 2020].

Mazlish, B. 2000. 'The man-machine and artificial intelligence'. In: Chrisley, R. ed. *Artificial intelligence: Critical concepts* (vol. 1). London: Routledge. 134 – 160.

McCarthy, J., Minsky, M.L., Rochester, N. & Shannon, CE. 1955. *A proposal for the Dartmouth summer research project on artificial intelligence*. [Conference proposal]. *Dartmouth College, Hannover, NH, August 31 1955*. Available from: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html> [Accessed on: 07 February 2020].

McClelland, R.T. 2016. Robotic alloparenting: a new solution to an old problem?. *The Journal of Mind and Behavior* 37(2): 71 – 98.

McKee, R. 1997. *Story: Substance, structure, style, and the principles of screenwriting*. New York, NY: ReganBooks.

McPhee, M., Baker, K.C. & Siemaszko, C. 2015. 'Deep Blue, IBM's supercomputer, defeats chess champion Garry Kasparov in 1997'. [Online]. *NY Daily News*. Available from: <https://www.nydailynews.com/news/world/kasparov-deep-blues-losingchess-champ-rooke-article-1.762264> [Accessed on: 02 March 2020].

Melon-Galvez, S. 2017. Intertextuality and the hedging system of the Filipino engineering students: Practices and pedagogy. *International Journal of Education* 9(2): 149 – 156.

Minor, J. 2014. '2004: A look back at the technology from 10 years ago'. [Online*]*. *Paste Magazine*. Available from: <https://www.pastemagazine.com/tech/tech-from-10-years-ago-blogging-bluetooth-and-the/> [Accessed on: 23 April 2020].

Monaco, J. 2009. *How to read a film: Movies, media, and beyond* (4th ed.). Oxford: Oxford University Press.

Moor, J. 2006. The Dartmouth College artificial intelligence conference: The next fifty years. *AI Magazine* 27(4): 87 – 89.

Moore, G.E. 1965. Cramming more components onto integrated circuits. *Electronics* 38(8): 114 – 117.

Morgan, M.S. and Wise, M.N. 2017. Narrative science and narrative knowing. Introduction to special issue on narrative science. *Studies in History and Philosophy of Science Part A* 62: 1 – 5.

Mori, H. 2012. 'The uncanny valley'. *IEEE Spectrum*. Available from: <https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley> [Accessed on: 06 April 2020].

Morning Consult. 2017. *National Tracking Poll 170401*. [Survey report]. Washington, DC: Morning Consult. Available from: <https://morningconsult.com/wp-

content/uploads/2017/04/170401_crosstabs_Brands_v3_AG.pdf> [Accessed on: 07 February 2020].

Nath, R. 2010. A Cartesian critique of the artificial intelligence. *Philosophical Papers and Reviews* 2(3): 27 – 33.

Newitz, A. 2016. 'Movie written by algorithm turns out to be hilarious and intense'. [Online]. *Ars Technica*. Available from: <https://arstechnica.com/the-multiverse/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/> [Accessed on: 02 May 2020].

Nieva, R. 2018. 'Alphabet chairman says Google Duplex passes Turing test in one specific way'. [Online]. *Cnet*. Available from: <https://www.cnet.com/news/alphabet-chairman-says-google-duplex-passes-turing-test-in-one-specific-way-io-2018/> [Accessed on: 03 March 2020].

Noble, D. 1999. *The religion of technology: The divinity of man and the spirit of invention*. New York, NY: Penguin Books.

Nomura, T., Suzuki, T., Kanda, T. & Kato, K. 2006. 'Altered attitudes of people toward robots: Investigation through the Negative Attitudes toward Robots Scale'. [Workshop proceedings]. *AAAI 2006 Workshop on Human Implications of Human-Robot Interaction, Boston, July 16 – 17 2006*. Menlo Park, CA: AAAI Press. 29 – 35.

Nourbhaksh, I.R. 2015. The coming robot dystopia: All too inhuman. *Foreign affairs (Council on Foreign Relations)* 94(4): 23 – 28.

Obozintsev, L. 2018. From Skynet to Siri: an exploration of the nature and effects of media coverage of artificial intelligence. Master of Arts in Communication. University of Delaware, Newark, DE.

Olivier, B. 2008. When robots would really be human simulacra: Love and the ethical in Spielberg's AI and Proyas's I, Robot. *Film-Philosophy* 12.2: 30 – 44.

Olivier, B. 2017. Artificial intelligence (AI) and being human: What is the difference? *Acta Academica* 49(1): 2 – 21.

Ophuls, W. 1977. *Ecology and the politics of scarcity: Prologue to a political theory of the steady state*. New York, NY: W. H. Freeman and Company.

Oppermann, A. 2018. 'Did Google Duplex beat the Turing test? Yes and no'. [Online]. *Towards Data Science*. Available from: < https://towardsdatascience.com/did-google-duplex-beat-the-turing-test-yes-and-no-a2b87d1c9f58> [Accessed on: 03 March 2020].

Orwell, G. 1949. *1984*. London: Secker & Warburg.

Pakdemirli, E. 2019. Perception of artificial intelligence (AI) among radiologists. *Acta Radiologica Open* 8(9) 1 – 2.

Palmer, R.B. 2008. 'Imagining the future, contemplating the past: The screen versions of *1984*'. In: Sanders, S.M. ed. *The philosophy of science fiction film*. Lexington, KY: The University Press of Kentucky. 171 – 190.

Parikh, P.C. 2019. AI film aesthetics: A construction of a new media identity for AI films. Master of Arts in Film Studies. Chapman University, Orange, CA.

Passchier, B. 2007. EVAM: Entertainment Value Assessment Matrix. *CILECT News*. Available from: <http://www.cilect.org/gallery/news/40/2007-11_hr.pdf> [Accessed on: 22 March 2020].

Peters, T. 2008. Allusions to theology: I, Robot, universalism and the limits of the law. *Media and Arts Law Review* 13: 77 – 92.

Pethokoukis, J. 2016. 'What the story of ATMs and bank tellers reveals about the 'rise of the robots' and jobs'. [Online]. *American Enterprise Institute*. Available from: <https://www.aei.org/economics/what-atms-bank-tellers-rise-robots-and-jobs/> [Accessed on: 20 April 2019].

Phua, C., Lee, V., Smith, K. & Gayler, R. 2012. A comprehensive survey of data mining-based fraud detection research. *Computers in Human Behaviour* 28(3): 1002 – 1013.

Plötz, T. and Fink, G.A. 2009. Markov models for offline handwriting recognition: A survey. *International Journal on Document Analysis and Recognition* 12(4): 269 – 298.

Porter, J.E. 1986. Intertextuality and the discourse community. *Rhetoric Review* 5(1): 34 – 47.

Powell, T.C. 2019. Can quantitative research solve social problems? Pragmatism and the ethics of social research. *Journal of Business Ethics (online)*. Available from: <https://link.springer.com/article/10.1007/s10551-019-04196-7> [Accessed on: 05 February 2020].

Prisecaru, P. 2016. Challenges of the fourth industrial revolution. *Knowledge Horizons – Economics* 8(1): 57 – 62.

Propp, V. 1968. *Morphology of the folktale* (2nd ed.). Austin, TX: University of Texas Press.

Rauch, S. 2018. Understanding the holocaust through film audience reception between preconceptions and media effects. *History and Memory* 30(1): 151 – 188.

Reinhart, R.J. 2018. *Most americans already using artificial intelligence products.* [Survey report]. Washington, DC: Gallup. Available from: <https://news.gallup.com/poll/228497/americans-already-using-artificial-intelligence-products.aspx> [Accessed on: 28 September 2019].

Ridley, M. 2010. *The rational optimist: How prosperity evolves*. HarperCollins e-books.

Righetti, N. and Carradore, M. 2019. From robots to social robots. Trends, representation and Facebook engagement of robot-related news stories published by Italian online news media. *Italian Sociological Review* 9(3): 431 – 454.

Ritchie, J., Lewis, J., McNaughton Nicholls, C. and Ormston, R. 2014. *Qualitative research methods* (2nd ed.). London: Sage.

Rohr, S. 2012. 'Pragmatism and semiotics'. In: Middeke, M., Müller, T., Wald, C. & Zapf, H. eds. *English and American Studies: Theory and Practice*. Stuttgart: Metzler. 220 – 224.

Rosenbaum, J. 2012. A matter of life and death: "A.I". *Film Quarterly* 65(3): 74 – 78.

Rossi, F. 2016. *Artificial intelligence: Potential benefits and ethical considerations: European parliament briefing.* PE 571.380. Available from: <http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_BRI(2016)571380> [Accessed on: 21 February 2020].

Rothstein, E. 2004. ‘Critic's notebook; For Asimov, robots were friends. Not so for Will Smith’. [Online]. *The New York Times*. Available from: <https://www.nytimes.com/2004/07/15/movies/critic-s-notebook-for-asimov-robots-were-friends-not-so-for-will-smith.html> [Accessed on: 12 April 2020].

Royal Society. 2018. *Portrayal and perceptions on AI and why they matter*. [Report]. London: Royal Society. Available from: <https://royalsociety.org/-/media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf> [Accessed on: 04 April 2019].

Sanders, S.M. 2008. ‘An introduction to the philosophy of science fiction film’. In: Sanders, S.M. ed. *The philosophy of science fiction film*. Lexington, KY: The University Press of Kentucky. 1 – 20.

Schatz, T. 1981. *Hollywood genres: Formulas, filmmaking, and the studio system*. Philadelphia, PA: Temple University Press.

Schwab, K. 2016. *The fourth industrial revolution: what it means, how to respond*. [Report]. World Economic Forum. Available from: <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/> [Accessed on: 01 April 2019].

Sculos, B.W. 2015. Automatons, robots, and capitalism in a very wrong twenty-first century: A review essay on Neill Blomkamp’s Chappie. *Class, Race and Corporate Power* 3(1): 1 – 7.

Searle, J.R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences* 3(3): 417 – 457.

Sebeok, T.A. 2001. *Signs: An introduction to semiotics* (2nd ed.). Toronto: University of Toronto Press.

Seed, D. 2011. *Science Fiction: A very short introduction*. Oxford: Oxford University Press.

Shelley, M. 1818. *Frankenstein*. London: Lackington, Hughes, Harding, Mavor & Jones.

Sjöberg, L. 2002. Attitudes toward technology and risk: Going beyond what is immediately given. *Policy Sciences* 35(4): 379 – 400.

Slocombe, W. 2016. ‘What science fiction tells us about our trouble with AI’. [Online]. *The Conversation*. Available from: <https://theconversation.com/what-science-fiction-tells-us-about-our-trouble-with-ai-58543> [Accessed on: 05 March 2020].

Smith, A. and Anderson, J. 2014. AI, robotics and the future of jobs. [Report]. *Pew Research Center*. Available from: <http://www.pewinternet.org/2014/08/06/future-of-jobs/> [Accessed on: 07 March 2020].

Smith, E. 2000. Thread structure: Rewriting the Hollywood formula. *Journal of Film and Video* 51(4): 88 – 96.

Smith, M. 1994. Character and emotional response in the cinema. *Cinema Journal* 33: 34 – 56.

Solomonoff, R.J. 1985. The time scale of artificial intelligence; Reflections on social effects. *Human Systems Management* (5): 149 – 153.

Spatola, N., Belletier, C., Normand, A., Chausse, P., Monceau, S., Augustinova, M., Barra, V., Huguet, P. & Ferrand, L. 2018. Not as bad as it seems: When the presence of a threatening humanoid robot improves human performance. *Science Robotics* 3(21): eaat5843.

Spiegeleire, S., Maas, M. & Sweijs, T. 2017. *Artificial intelligence and the future of defense: strategic implications for small and medium-sized force providers*. [Report]. The Hague: The Hague Centre for Strategic Studies.

Staiger, J. 2005. *Media reception studies*. New York, NY: NYU Press.

States, B.O. 1985. The anatomy of dramatic character. *Theatre Journal* 37(1): 86 – 101.

Stoehr, K.L. 2008. ‘2001: A philosophical odyssey’. In: Sanders, S.M. ed. *The philosophy of science fiction film*. Lexington, KY: The University Press of Kentucky. 119 – 134.

Stokes, J. 2013. *How to do culture and media studies*. London: Sage.

Stone, Z. 2017. ‘Everything you need to know about Sophia, the world's first robot citizen’. [Online]. *Forbes*. Available from: <https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#5934e5846fa1> [Accessed on: 18 January 2020].

Swant, M. 2017. 'As IBM ramps up its AI-powered advertising, can Watson crack the code of digital marketing?'. [Online]. *Advertising Week*. Available from: <https://www.adweek.com/digital/ibm-is-bringing-next-level-ai-technology-to-marketers/> [Accessed on: 20 February 2020].

Taysina, E. 2013. Running head: Social semiotics: Unaccomplished project. *Procedia - Social and Behavioral Sciences* 92: 907 – 914.

Telotte, J.P. 2001. *Science fiction film*. Cambridge: Cambridge University Press.

Todorov, T. 1977. *The Poetics of Prose*. Ithaca, NY: Cornell University Press.

Tomaselli, K.G. 1996. *Appropriating images: The semiotics of visual representation*. Højbjerg: Intervention Press.

Truitt, B. 2015. Robots storm the cinema: From 'Chappie' to 'Ultron' to ' Terminator', marvels of machinery spotlight dreams and nightmares of artificial intelligence. [Online]. USA Today. Available from: <https://www.pressreader.com/usa/usa-today-us-edition/20150304/282325383433837> [Accessed on 01 August 2020].

Turing, A. 1950. Computing machinery and intelligence. *Mind* 59(236): 433 – 460.

Turkle, S. 2011. *Alone together: Why we expect more from technology and less from each other*. New York, NY: Basic Books.

UK House of Lords. 2018. *AI in the UK: ready, willing and able? Select committee on artificial intelligence, report of session 2017 – 19*. [Report]. HL Paper 100. London: Dandy Booksellers Limited.

Van Leeuwen, T. 2005. *Introducing social semiotics*. London: Routledge.

Van Melle, W. 1978. MYCIN: a knowledge-based consultation program for infectious disease diagnosis. *International Journal of Man-Machine Studies* 10(3): 313 – 322.

Vannini, P. 2007. Social semiotics and fieldwork: Method and analytics. *Qualitative Inquiry* 13(1): 113 – 140.

Walliman, N. 2011. *Research methods: The basics*. London: Sage.

Webb, J. 2009. *Understanding representation*. London: Sage.

White, V.A. 2008. 'A.I.: Artificial Intelligence: Artistic indulgence or advanced inquiry?'. In: Kowalski, D.A. ed. *Steven Spielberg and philosophy: We're gonna need a bigger book*. Lexington, KY: The University Press of Kentucky. 210 – 226.

Wikes, M.V. 1956. *Automatic digital computers*. New York, NY: John Wiley & Sons.

Wilson, A. 2017. Techno-optimism and rational superstition. *Techné: Research in Philosophy and Technology* 21(2): 1 – 21.

Xu, M., David, J.M. and Kim, S.H. 2018. The fourth industrial revolution: Opportunities and challenges. *International Journal of Financial Research* 9(2): 90 – 95.

Zabarah, R. 2015. 'Neill Blomkamp: Artificial intelligence will do what it wants'. [Online]. *Popular Mechanics*. Available from: <https://www.popularmechanics.com/culture/movies/interviews/a14365/neill-blomkampinterview-chappie/> [Accessed on: 14 May 2020].

Zhang, B. and Dafoe, A. 2019. *Artificial intelligence: American attitudes and trends*. [Survey report]. Oxford: University of Oxford Press. Available from: <http://tiny.cc/aipublic2019pdf> [Accessed on: 12 February 2019].

ADDENDUM A: AUDIENCE RESEARCH QUESTIONNAIRE

Section 1: Eligibility check and demographic information

1.1) Did you watch the film [*I, Robot* (2004) or *Chappie* (2015)] between 2010 and 2020, and are you over 18 years old?

a) Yes

b) No

1.2) Please specify where you viewed the film

a) Cinema

b) DVD / Blu-Ray rental

c) Television

d) Online streaming platform

e) Other (please specify)

1.3) How old are you?

a) 18 – 25

b) 25 – 30

c) 30 – 35

d) 35 – 40

e) over 40

1.4) What is your gender?

a) Male

b) Female

c) Non-binary

d) Prefer not to say

1.5) What is your highest level of education?

a) Some High School

b) High School Graduate

c) Post-secondary degree / diploma / certificate

d) Post-graduate degree

e) Other (please specify)

f) Prefer not to say

1.6) How would you categorise your employment industry?

a) Agriculture, Food and Natural Resources

b) Architecture and Construction

c) Arts, Audio/Video Technology & Communications

d) Business Management & Administration

e) Computer science / information technology

f) Education and training

g) Finance

h) Government & Public Administration

i) Health Science / services

j) Hospitality & Tourism

k) Law, Public Safety, Corrections & Security

l) Manufacturing

m) Marketing, Sales and Service

n) Science, Technology, Engineering & Mathematics

o) Transportation, Distribution & Logistics

p) Student (please specify degree)

q) Currently Unemployed

r) Other (please specify)

1.7) Are you religious? If so, please specify your religion

a) Yes

b) No

c) Prefer not to say

1.8) What is your nationality?

1.9) Please select all applications / services that you use regularly from the list below

a) Search engines

b) Navigation software

c) Translation software

d) Digital personal assistants

e) Music or video streaming services

f) Facebook, Snapchat or Reddit

g) Amazon or Ebay

Section 2: Impressions of the portrayal of strong humanoid AI

How much do you agree with the following statements?

2.1) The film was engaging / entertaining

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

2.2) The artificially intelligent humanoid robots portrayed in the film were portrayed in a positive light

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

2.3) The artificially intelligent humanoid robots portrayed in the film were portrayed more positively than most other science fiction films I have watched

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

2.4) I cared about one or more of the artificially intelligent humanoid robots portrayed in the film

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

2.5) The film helped increase my understanding about artificial intelligence and consciousness

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

2.6) I felt more positive about future artificial intelligence after viewing the film than before viewing it

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

2.7) I have, at least once, thought about the film when seeing, hearing, or saying something else about artificial intelligence

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

2.8) Artificially intelligent humanoid robots will be created that behave in the same way as the robots portrayed in the film

a) Strongly disagree
b) Disagree
c) Neither agree nor disagree
d) Agree
e) Strongly agree

Optional: Provide reasons for your answer

2.9) I'd like to live in a world in which artificially intelligent humanoid robots, such as those portrayed in the film, exist

a) Strongly disagree
b) Disagree
c) Neither agree nor disagree
d) Agree
e) Strongly agree

Optional: Provide reasons for your answer

2.10) The film changed my opinion of the possibility of conscious artificially intelligent humanoid robots being created

a) Strongly disagree
b) Disagree
c) Neither agree nor disagree
d) Agree
e) Strongly agree

Optional: Provide reasons for your answer

Section 3: General attitudes towards artificial intelligence

How much do you agree with the following statements?

3.1) I often interact with services or applications which make use of artificial intelligence

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

3.2) I know about the trends and use of current and upcoming technology

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

3.3) I am optimistic about future technologies involving artificial intelligence

a) Strongly disagree

b) Disagree

c) Neither agree nor disagree

d) Agree

e) Strongly agree

Optional: Provide reasons for your answer

3.4) Which statement best describes your feelings towards future developments to artificial intelligence?

a) I believe that it will ultimately be disastrous/destructive for humans and/or the planet

b) I believe that it will ultimately be bad for humans and/or the planet

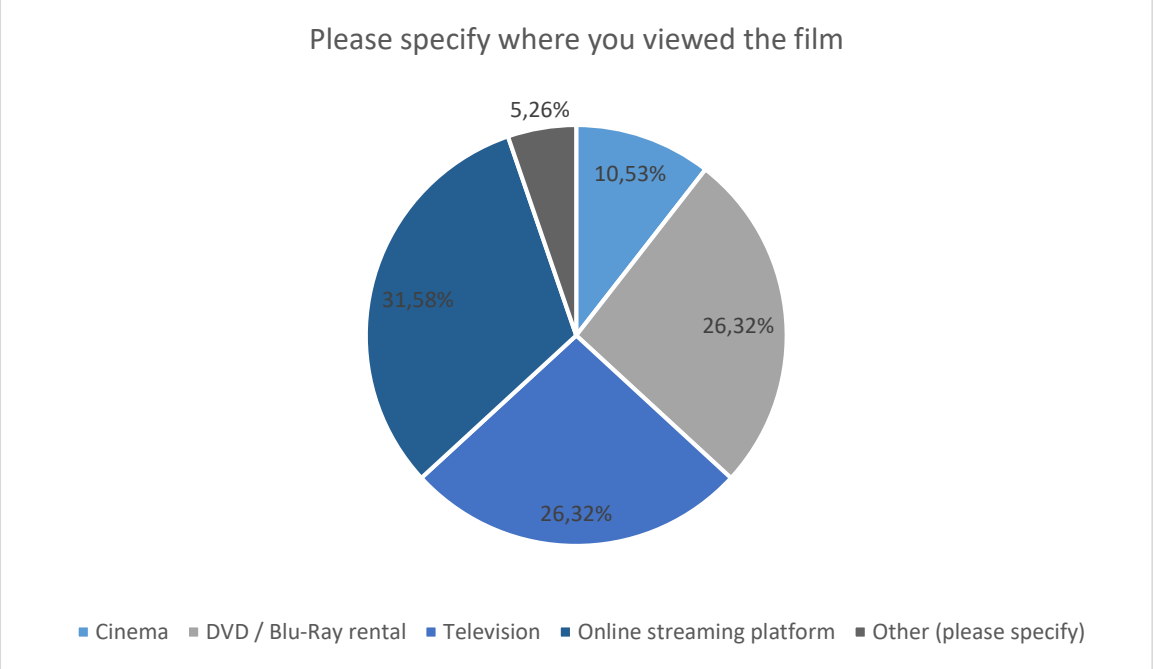
c) I believe that it will neither be good or bad, and will not significantly change society

d) I believe that it will ultimately be good for humans and/or the planet

e) I believe that it will ultimately be great for humans and/or the planet

Optional: Provide reasons for your answer

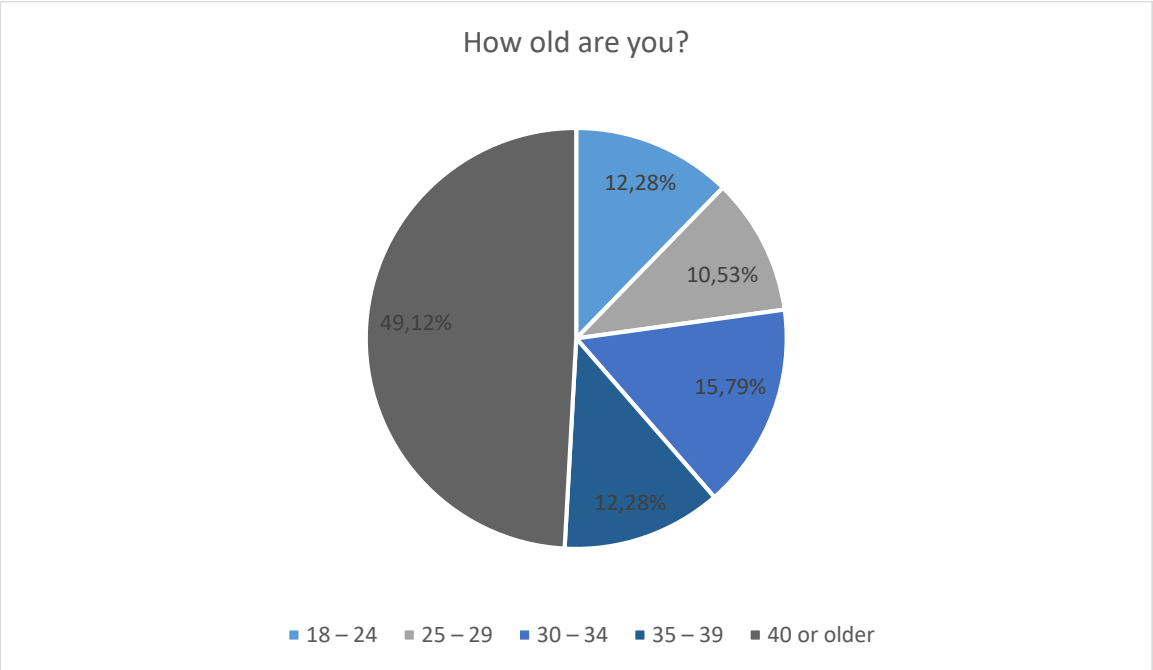
ADDENDUM B: DEMOGRAPHIC DATA FOR *I, ROBOT* (2004)



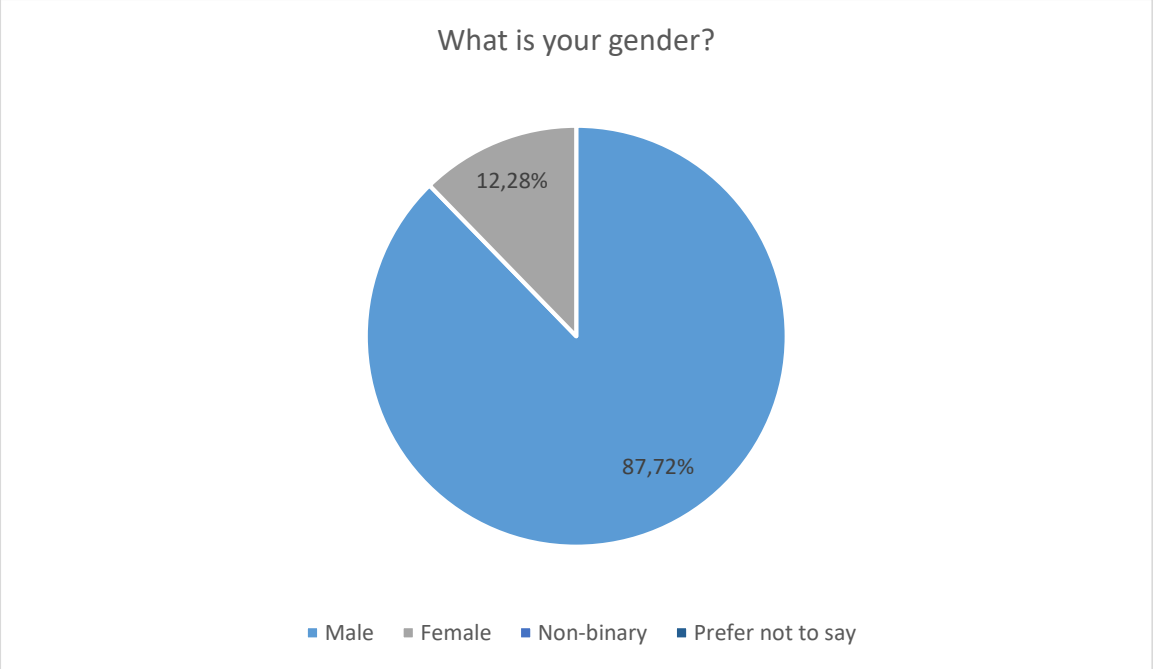
***I, Robot* (2004) survey responses for question 1.2**

The 5,26% of respondents selecting 'other' for question 1.2 included the following quantitative information:

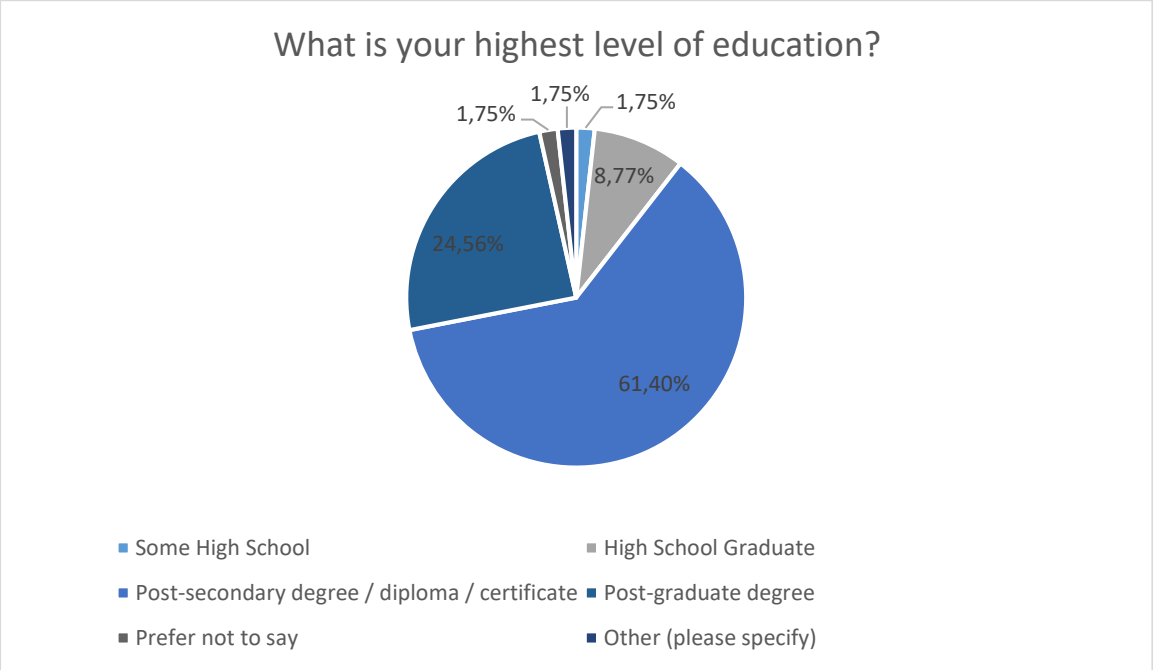
1) Cinema upon release, rewatched on DVD
2) Projector for a College Class
3) Cinema first, DVD second. There were no theatrical options for 2010-2020



***I, Robot* (2004) survey responses for question 1.3**



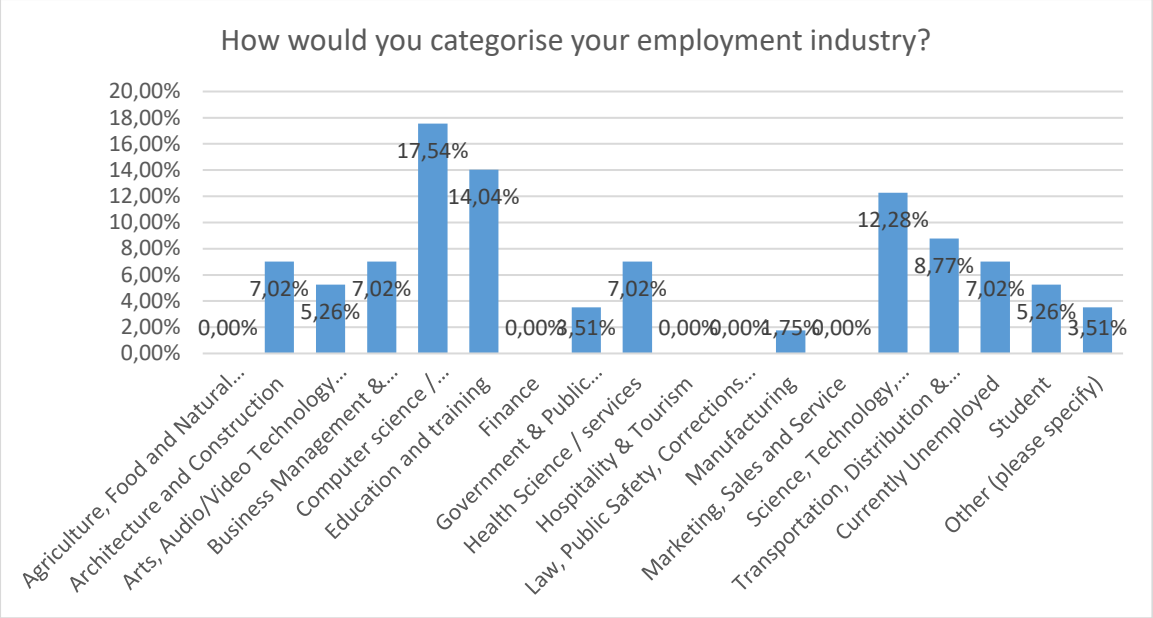
***I, Robot* (2004) survey responses for question 1.4**



***I, Robot* (2004) survey responses for question 1.5**

The only respondent selecting the 'other' category for question 1.5 wrote the following qualitative information:

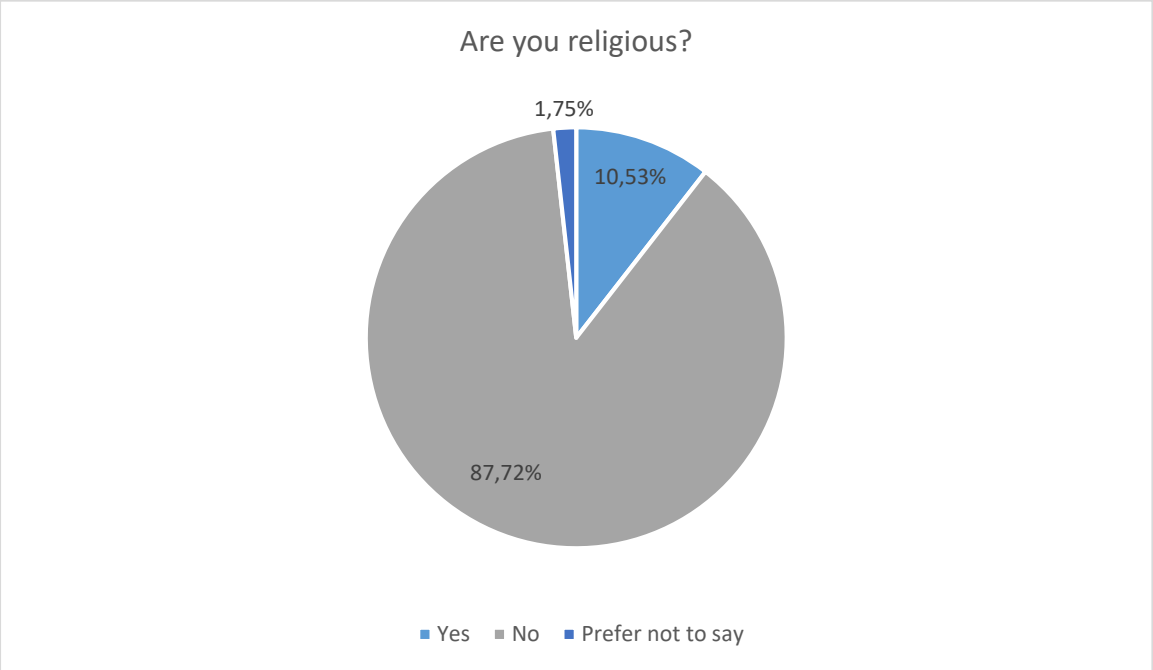
1) Some college



***I, Robot* (2004) survey responses for question 1.6**

The 3,51% of respondents selecting 'other' for question 1.6 included the following qualitative information:

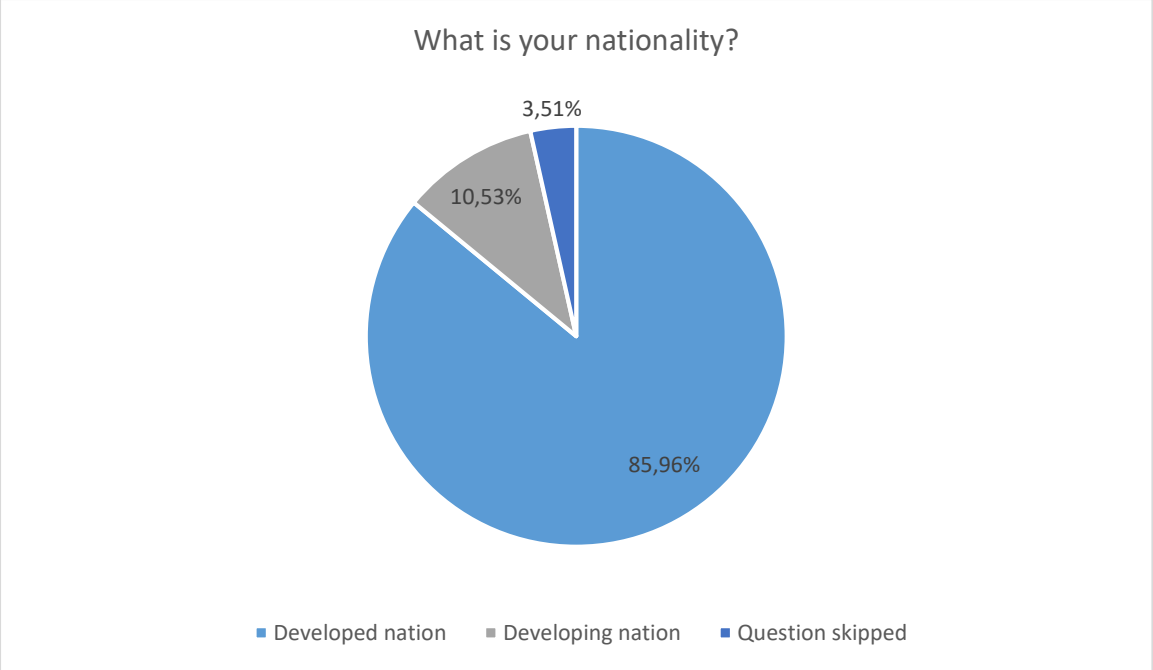
1) Hd mechanic
2) Homemaker



***I, Robot* (2004) survey responses for question 1.7**

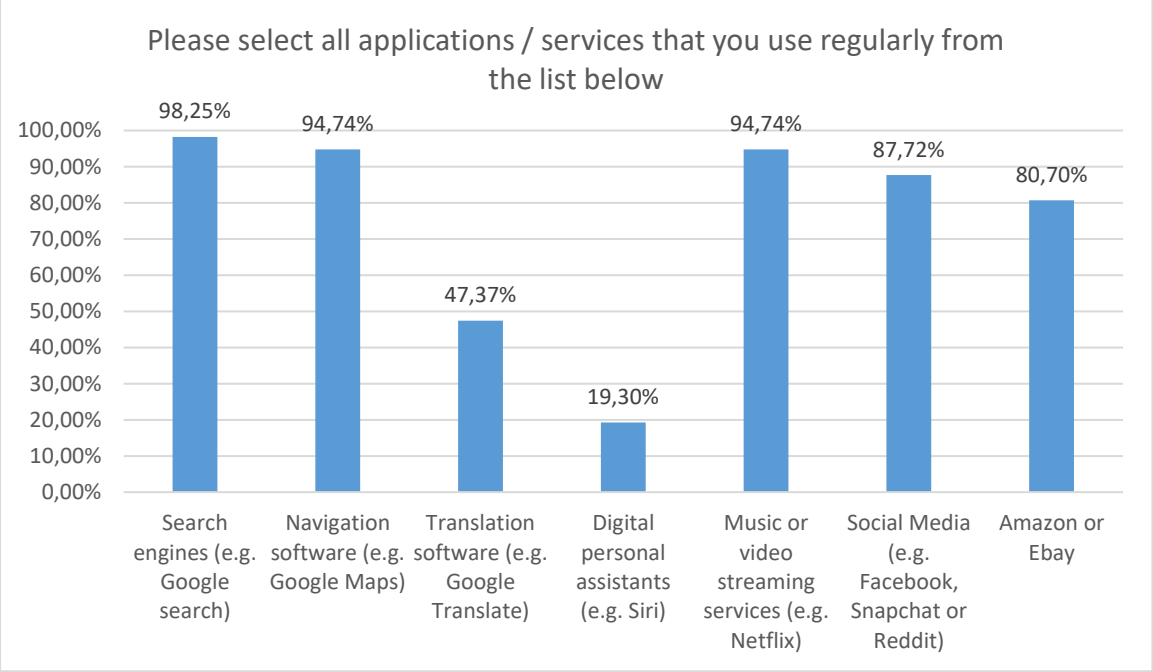
The 10,53% of respondents selecting 'yes' for question 1.7 included the following qualitative information:

1) Follows the Christian religion but not religious
2) Christian/spiritual
3) orthodox Christian
4) Christian: Evangelical, Grace-based
5) Orthodox Christian



***I, Robot* (2004) survey responses for question 1.8**

The respondents' nationality for question 1.8 was organised manually into the developed or developing nation category after the fact based on the qualitative response provided.



***I, Robot* (2004) survey responses for question 1.9**

ADDENDUM C: QUALITATIVE SURVEY RESPONSES FOR *I, ROBOT* (2004)

The film was engaging / entertaining	
Response Category	Comment
Neither agree nor disagree	1. Mostly forgettable; lots of product placement 2. having read the book i was disappointed in the divergence.
Agree	1. It's my kind of trash. I hate it and love it. Shia is awesome. 2. It was a nice looking movie for the time 3. Based on interesting concepts (laws of robotic) 4. I am a longtime fan of the original author Isaac Asimov and was excited for this films release. 5. It was fun but not true to the book wich is more cerebral 6. The plot was meh. 7. Big fan of the short stories that the movie was based, was expecting a different story but still enjoyed 8. Clever writing and good choice in actors It was a studio adaptation cincentrating more on blockbuster aesthetics than the point of the original material.
Strongly agree	1. Mystery, relationships, acting, pace, cinematography 2. I tend to think of it as more of an action movie rather than anything especially philosophically engaging, but it’s still a great watch every time

Qualitative responses for *I, Robot* (2004) question 2.1

The artificially intelligent humanoid robots portrayed in the film were portrayed in a positive light	
Response Category	Comment
Strongly disagree	1. Murder bad 2. They were mostly villains. Why would anyone design them to look mad and change color when reprogrammed?

Disagree	1. AI was portrayed as trying to install a totalitarian regime upon humanity. 2. Plot of film was based on if robots were OK or not. Protaganist was anti-robot which became prevailing storyline.
Neither agree nor disagree	3. One was positive, the rest negative. 4. The AI characters take opposite roles - the main character being good, the enemy AI being bad. 5. AI in the film were both benign and hostile 6. Other than Sonny, all the robots were relatively at the whim of the people in charge of them, reprogramming ect. Humans trying to influence the perception of the robots tarnished the image of the robots. 7. As in the novel, the robots we're both the ultimate protagonist and antagonist 8. It does both. They're mostly bad, but I also feel like there's an implied potential for good through "Sonny" the robots only seem bad because they're essentially slaves, but they can evolve "one day they'll have feelings, one day they'll have dreams." 9. There are good (sunny) and bad (Viki) representations of AI. But the humanoid AI besides sunny don't have any autonomy really to be able to asses how they are being portrayed. They are mainly shown as being a mere tool for humans to use in most cases.
Agree	1. Apart from the evil computer the AI was friendly
Strongly agree	3. There's only one artificially intelligent humanoid robot in the film, and it's hard to argue that Sonny is portrayed positively. All the other robots aren't AI.

Qualitative responses for *I, Robot* (2004) question 2.2

The artificially intelligent humanoid robots portrayed in the film were portrayed more positively than most other science fiction films I have watched	
Response Category	Comment
Strongly disagree	1. On average robots are portrayed badly however these were robots taking over the world the only worse ones that come to mind are skynet from terminator
Disagree	1. Maybe compared to skynet, but not compared to the average AI use. 2. Most films I've seen have put the robots either in the role of pure badguy or as sidekick/protector with a hint of comedic relief. I Robot was more complex and though I didn't think it went as far the inspiring books and was hindered by its "action movie" status, I do think it touched on the more complex idea that robots actions will be the result of how humans choose to make them. But that not every situation can be anticipated and programmed for and so there may be surprising emergent behavior in the robots, saving Spooner instead of the child, deciding that humanity must be protected from itself. Odd emergent behaviour and humans interpreting these as glitches was a running theme in Asimov's works. 3. The plot required a negative/neutral attitude towards AI robots. 4. As stated previously none have any real autonomy and for ease of mass appeal they are shown to be flawed and easily stopped / outsmarted which I do not believe is the case.
Agree	1. Of the main character, there is time spent exploring feelings and dreams, making the character seem likeable and having a personality.

Strongly agree	1. AI is almost always the bad guy, and Sonny is the hero.

Qualitative responses for *I, Robot* (2004) question 2.3

I cared about one or more of the artificially intelligent humanoid robots portrayed in the film	
Response Category	Comment
Agree	1. The main robot protagonist was warm and engaging
Strongly agree	1. Again, there is only one, and he's the most sympathetic person in the film. 2. Sonny was well done, great animation, great voice actor, great scripting, showed empathy. 3. One. Sunny.

Qualitative responses for *I, Robot* (2004) question 2.4

The film helped increase my understanding about artificial intelligence and consciousness	
Response Category	Comment
Strongly disagree	1. Abso-goddamn-lutely not. This film is hot garbage. 2. It's an extremist version of AI, not remotely linked to any actual possibilities
Neither agree nor disagree	1. I was probably already a little above the average knowledge base but the film did give some interesting ideas about a robot integrated society and its problems. 2. It introduces the ideas but never really explores them in great depth, just brushes the surface so as not to put of casual movie goers.
Agree	1. First time I'd heard Asimov's Three Laws of AI - interesting lead in to how AI could be created in a way that was safe to humans
Strongly agree	1. This brought up ideas/notions that were unique to me.

Qualitative responses for *I, Robot* (2004) question 2.5

I felt more positive about future artificial intelligence after viewing the film than before viewing it	
Response Category	Comment
Disagree	1. I thought it was overall negative about the risks of A.I. to the world. Sonny was extremely positive but it portrayed an attempted A.I. coup 2. The whole story is about unforeseen consequences and loopholes being dangerous. Trial and error phase of implementing artificial intelligence in a way that was depicted in the movie is bound to be a dangerous time 3. The film was designed to portray AI as easily manipulated and potentially dangerous. It did not accurately portray the content of the source material.
Neither agree nor disagree	1. I did not feel positive, I left the film feeling informed and careful. 2. Again hard to assess, due to lack of any real exploration of idea. Mainly just looks at exploiting technophobic stereotypes of what AI could and might be

Qualitative responses for *I, Robot* (2004) question 2.6

I have, at least once, thought about the film when seeing, hearing, or saying something else about artificial intelligence	
Response Category	Comment
Strongly disagree	1. I think about this film in the context of Will Smith dual-wielding guns on a motorbike. 2. Only by it sharing the same name as the book
Agree	1. When this topic comes up people are want to mention I robot as a critical example 2. The premise of giving away freedom (to a save keeper here: the ai) for safety/ protection is more present than ever

	3. i understand that those who do not know AI or Asimov would be misled by the movie.
Strongly agree	1. This film brought forward profound ideas on humanity, good/bad. Unique progressive ideas and concepts.

Qualitative responses for *I, Robot* (2004) question 2.7

Artificially intelligent humanoid robots will be created that behave in the same way as the robots portrayed in the film	
Response Category	Comment
Strongly disagree	1. Is this a serious question?
Disagree	1. Humanoid robots are too close to us I dont believe people would engage with the in the way portrayed in I robot they would not be treated as a memebr of the family akin to a dog more as a servent thus making them humanoid would be pointless 2. I don't believe AI will have personalities such as those shown in the film 3. Probably not while I'm still alive. 4. It doesn't seem likely but that doesn't mean I'm right
Neither agree nor disagree	1. Might be the case. Hard to say specifically. 2. I do think the 3 laws are a good idea.
Agree	1. It seems the logical progression. Even though non-humanoid robots have a presence in the world today they are not (to my knowledge) sentient. I feel sentient humanoid robots will allow us to relate to them better so we can tolerate them and grow to rely on them.
Strongly agree	1. The robots yes. Maybe not the AI. Need to ensure the two aren't conflated.

Qualitative responses for *I, Robot* (2004) question 2.8

I'd like to live in a world in which artificially intelligent humanoid robots, such as those portrayed in the film, exist	
Response Category	Comment
Strongly disagree	1. The entire premis of the film resolves around the fact that the Robots portrayed in the movie are to a degree self aware, using them for mundane taks is slavery
Disagree	1. I worry what will become of humanity if we create an effective slave race of self-aware robots.
Neither agree nor disagree	1. I feel insecure about the robots being smarter, stronger, faster, and longer-lived than me.
Agree	1. I do think they are cool I would like one 2. Assuming the Nestor 5s don't take over then yes. 3. Just not the killer ones ok?
Strongly agree	1. I would also like to live in Middle Earth or on the USS Enterprise.

Qualitative responses for *I, Robot* (2004) question 2.9

The film changed my opinion of the possibility of conscious Artificially Intelligent humanoid robots being created	
Response Category	Comment
Strongly disagree	1. The film makes no effort to go into the logistics of creating AI.
Neither agree nor disagree	1. The film did not change my opinion. I have been reading scifi stories for years that deal/dealt with robots.

Qualitative responses for *I, Robot* (2004) question 2.10

I often interact with services or applications which make use of Artificial Intelligence	
Response Category	Comment
Neither agree nor disagree	1. The spectrum of AI is huge, from calculators to a singularity event, this definition is too broad to answer. 2. I suppose if the AI are programmed good enough, I would not be able to tell the difference robot/person.
Agree	1. Not real AI but has smart behaviours

Strongly agree	1. any time I use google there is an AI in there.

Qualitative responses for *I, Robot* (2004) question 3.1

I know about the trends and use of current and upcoming technology	
Response Category	Comment
Disagree	1. I know a little about current tech, but I'm probably way behind on future tech.
Agree	1. * In the fields that concern me directly I am aware. No single person can possibly know of all new advancements and trends. 2. I'm not on the edge but I keep up. 3. I read alot

Qualitative responses for *I, Robot* (2004) question 3.2

I am optimistic about future technologies involving Artificial Intelligence	
Response Category	Comment
Disagree	1. I think it has amazing potential but is too open to abuse from the humans who control/create it.
Neither agree nor disagree	1. The spectrum of AI is huge, from calculators to a singularity event, this definition is too broad to answer. 2. Depends entirely on the way a.i. is used 3. I am not really optimistic about this. Neutral is a better descriptor. 4. I would be, but mankind is famously irresponsible with this type of thing
Agree	1. Useful pointy stick
Strongly agree	1. Just like anything. If used properly it can accomplish great things. Good or bad.

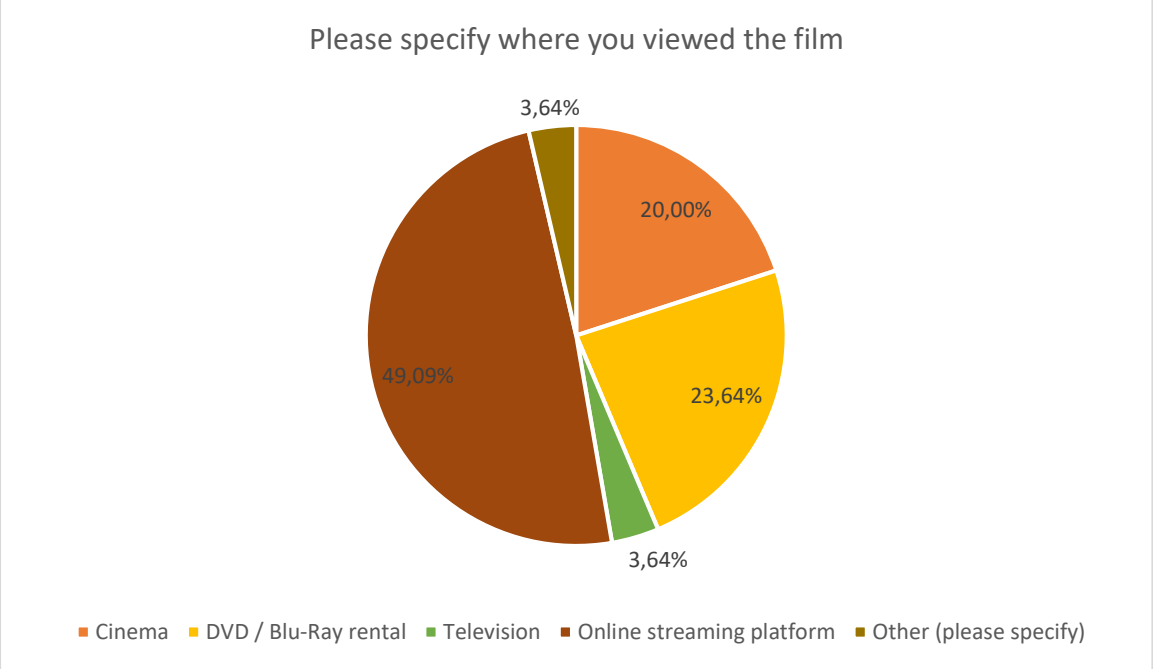
Qualitative responses for *I, Robot* (2004) question 3.3

Which statement best describes your feelings towards future developments to artificial intelligence?	
Response Category	Comment
I believe that it will ultimately be bad for humans and/or the planet	1. I think it could be great but it unlikely we will use it for good. 2. I think it will probably be good for the planet, bad for humans. But like in the

	movie it will probably be good for a while, before it gets bad
I believe that it will neither be good or bad, and will not significantly change society	1. It could go either way, but I don't "believe" any of these answers. I can hope for the latter, but hopes and dreams are not beliefs. 2. I think it will be a significant change, but am not confident whether it will be overall positive or negative 3. As long as we program them carefully, they will not change society, they will enhance it. 4. I have a hard time believing that it will be only good or only bad. Things like this are momentous. It's what people do with it an individual basis that is going to make the big impact
I believe that it will ultimately be good for humans and/or the planet	1. We screw up everything. I just hope for a net positive.
I believe that it will ultimately be great for humans and/or the planet	1. An AI assigned to something like say, resource management, would be able to do that in a totally unbiased and incorruptible manner. 2. Emphasis on the planet
No response	1. I honestly dont know because I think its stupid to presume Robots will think like us, I imagine they'll try to model them in our image but I dont think its possible to make a digital brain that will work exactly like our biological one. So for it's me its impossible to what's will come from AI

Qualitative responses for *I, Robot* (2004) question 3.4

ADDENDUM D: DEMOGRAPHIC DATA FOR *CHAPPIE* (2015)

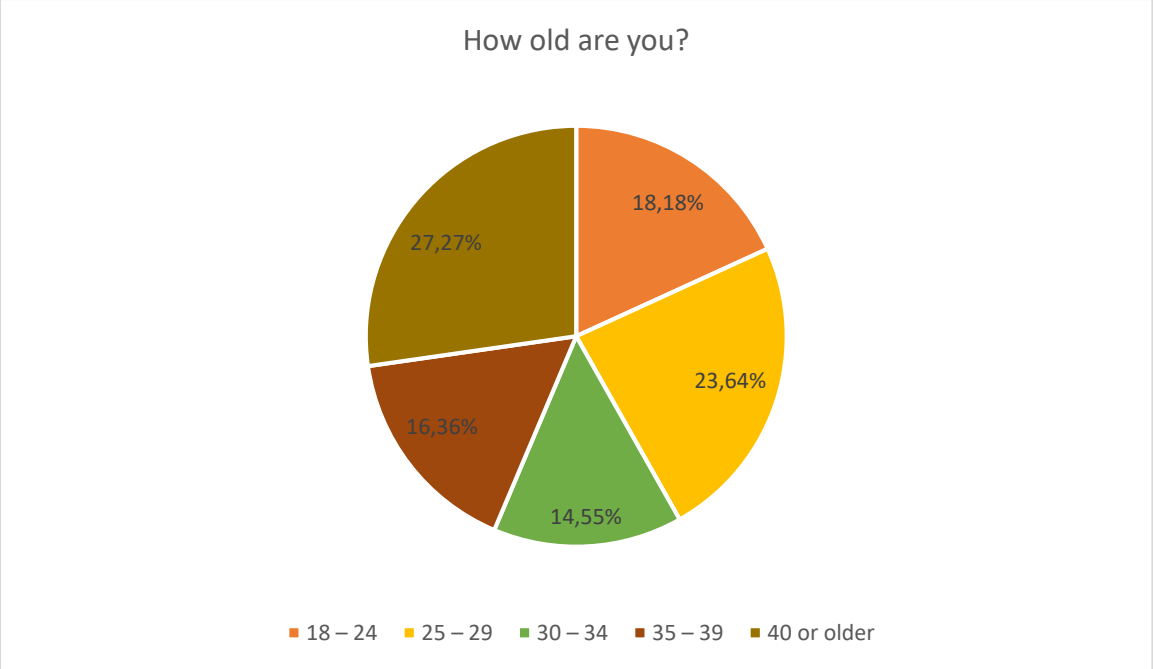


***Chappie* (2015) survey responses for question 1.2**

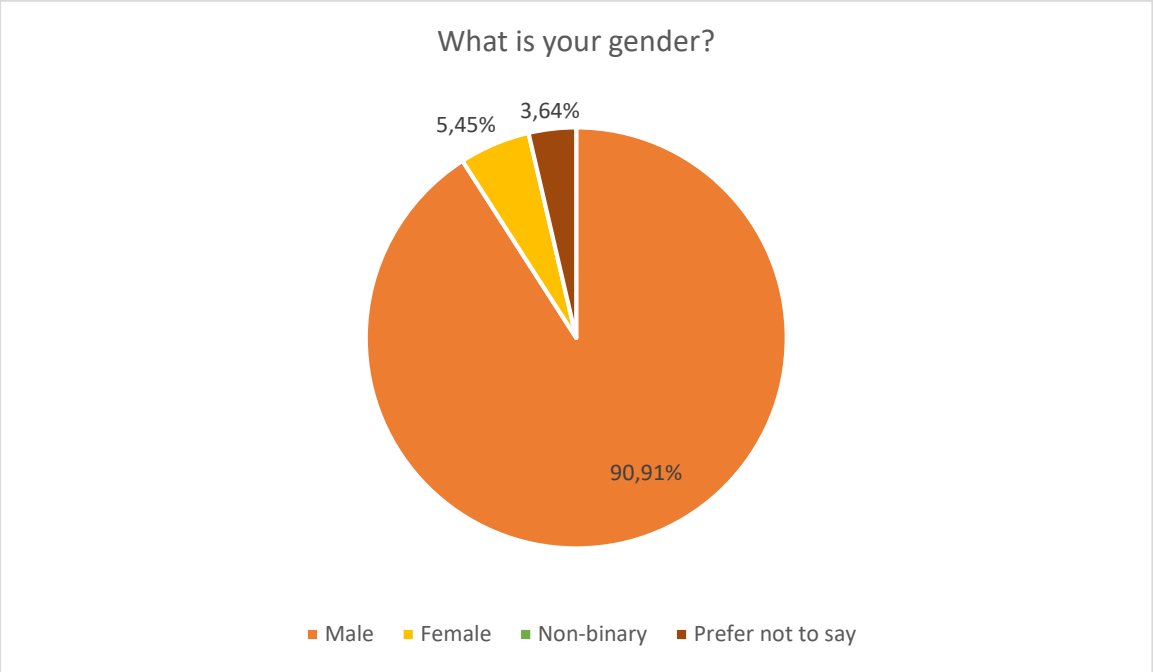
The 3,64% of respondents selecting 'other' for question 1.2 included the following quantitative information:

1) Torrent 1080p dts Blu-ray rip
2) mkv download

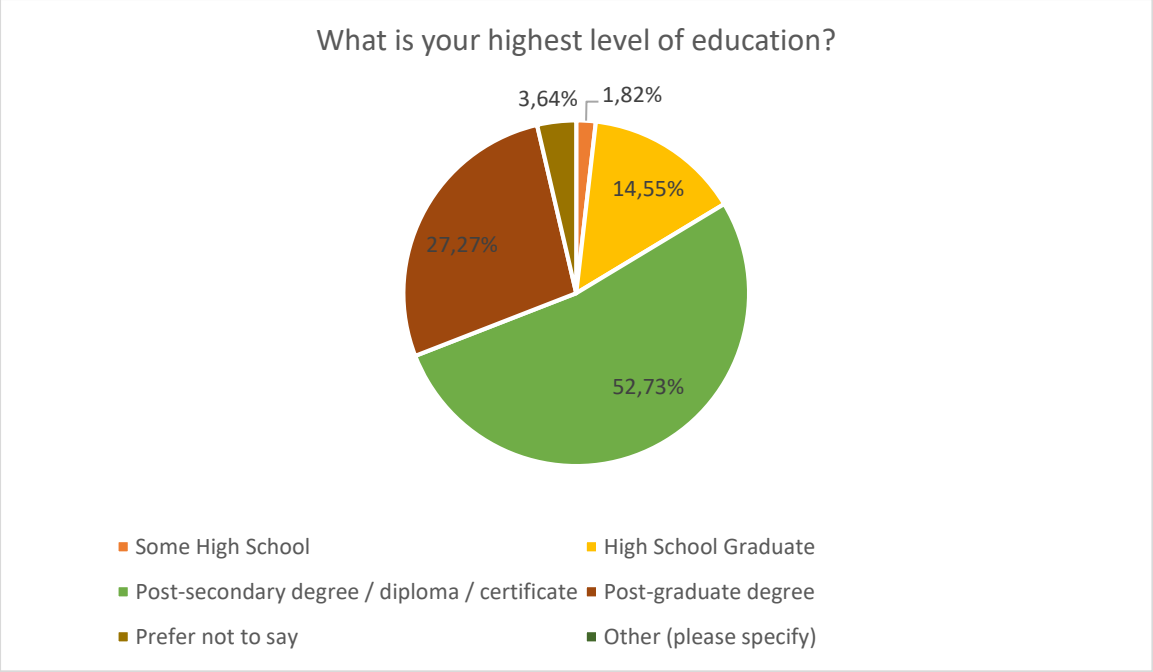
This information suggests that the material was illegally downloaded or shared.



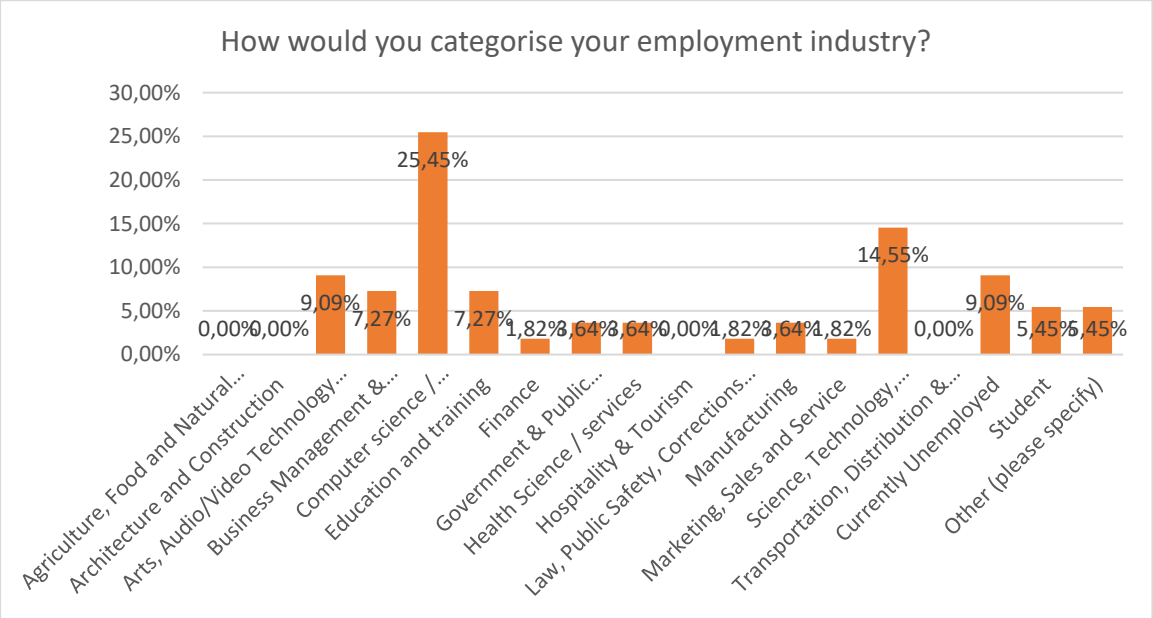
***Chappie* (2015) survey responses for question 1.3**



***Chappie* (2015) survey responses for question 1.4**



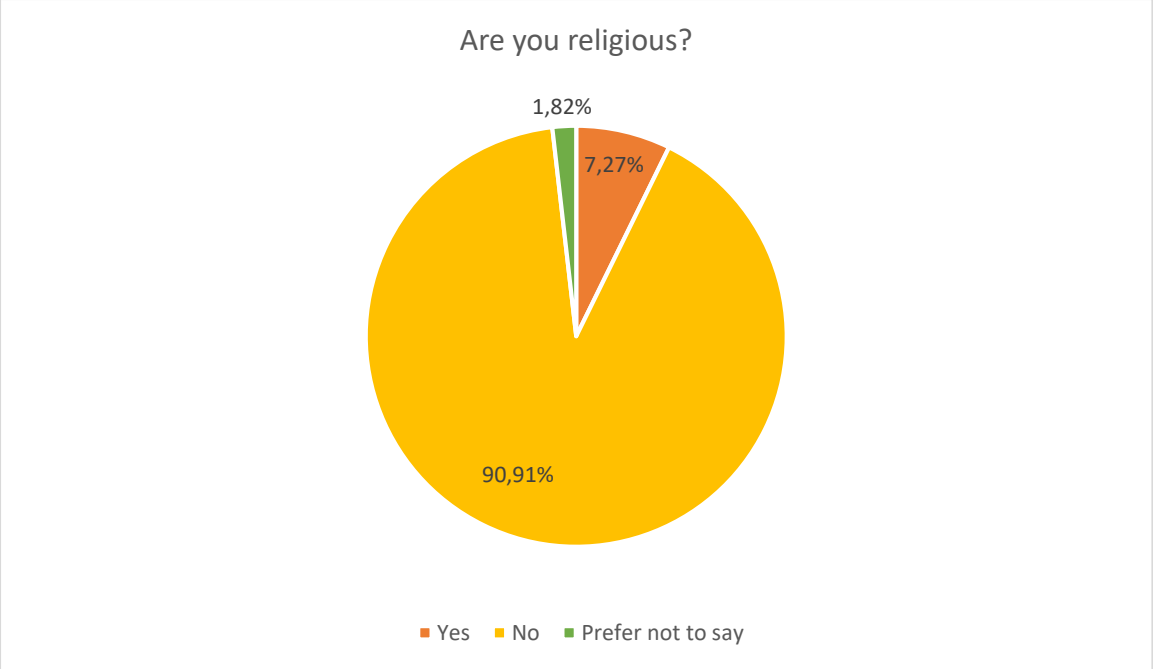
***Chappie* (2015) survey responses for question 1.5**



***Chappie* (2015) survey responses for question 1.6**

The 5,45% of respondents selecting 'other' for question 1.6 included the following qualitative information:

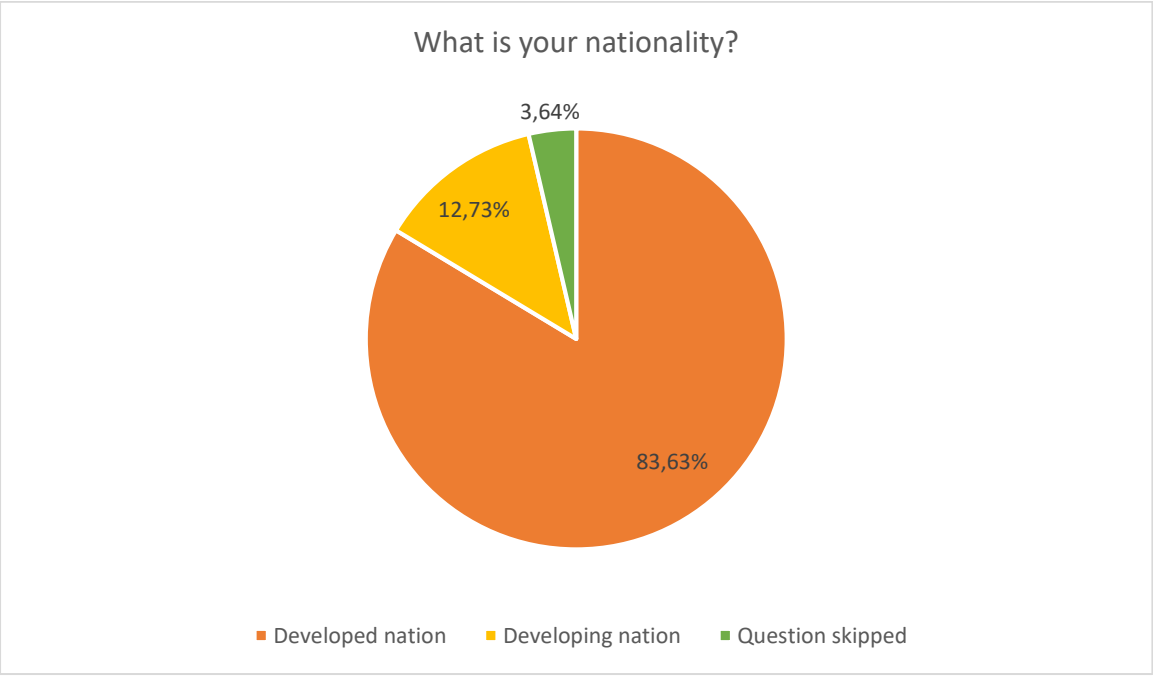
1) Homemaker
2) retired



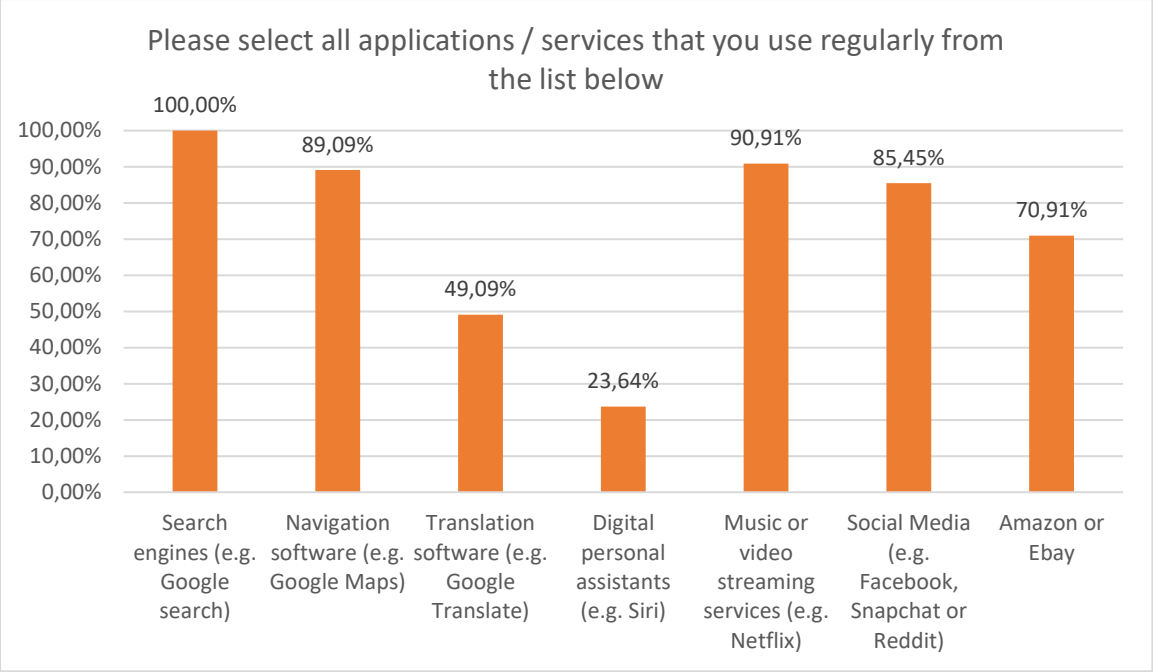
***Chappie* (2015) survey responses for question 1.7**

The 7,27% of respondents selecting 'yes' for question 1.7 included the following qualitative information:

1) Protestant Christian
2) Christian



***Chappie* (2015) survey responses for question 1.8**



***Chappie* (2015) survey responses for question 1.9**

ADDENDUM E: QUALITATIVE SURVEY RESPONSES FOR *CHAPPIE* (2015)

The film was engaging / entertaining	
Response Category	Comment
Strongly disagree	1. A unique outlook on AI portrayed as an innocence (it's been done, but not nearly explored enough).
Disagree	1. Some unrelatable and annoying characters (Die Antwoord) and cartoonish villian.
Neither agree nor disagree	1. There were some interesting bits but for the most part the story and most character motivations were all over the place
Agree	1. Great premise, poorly executed. I love Die Antwoord, but they ruined the movie. 2. I think I wanted to like it more than I actually liked it 3. Die antword were a strange choice but I like them so thumbs up from me
Strongly agree	1. I enjoyed seeing Chappie getting molded by his sorroundings and sorrounding people 2. The film featured one of the best mo-cap performances i have ever seen. "We did a bunch of tests before we started shooting where I just kind of improvised the scenes with another actor, to work out how Chappie might behave at different ages," says Copley. It paid of: the robot nails an arc that most actors would die for, and our hearts along the way. 3. I love exploring different worlds, that are only possible in science-fiction

Qualitative responses for *Chappie* (2015) question 2.1

The artificially intelligent humanoid robots portrayed in the film were portrayed in a positive light	
Response Category	Comment
Neither agree nor disagree	1. Most were negative, Chappie was portrayed in a positive light. 2. I think they showed the good and the bad sides very well. How Chappie helps

	commiting crimes because he doesnt know better, but also saving his maker and protecting those which he cares about.
Agree	1. I think Chappie was portrayed positively, but not the other robots. 2. Chappie was a relatable and good lad reminds me of Johnny 5 3. The movie was mostly on the nurture side of the nature v nurture debate - Chappie shoots guns, swears and commits crime but those are all things he learned from 3 out of 4 of his primary caregivers. He also tried to be kind and so good because those are the values the 4th caregivers tries to instill in him.
Strongly agree	1. Everything he thought or did was taught. So if we teach AI to be good they will be

Qualitative responses for *Chappie* (2015) question 2.2

The artificially intelligent humanoid robots portrayed in the film were portrayed more positively than most other science fiction films I have watched	
Response Category	Comment
Disagree	1. I think its important to consider context of the world which the AI lives in. So for me its hard to compare
Agree	1. Their purpose was negative but that's down to the powerhungry people and corporations in control of them. 2. Chappie was portrayed positively with the scout robots in general also considered good even the big bad robot was shown as mindless controlled by the villian character

Qualitative responses for *Chappie* (2015) question 2.3

I cared about one or more of the artificially intelligent humanoid robots portrayed in the film	
Response Category	Comment
Neither agree nor disagree	1. I found the movie's attempts to make me care about Chappie a bit too on the nose which annoyed me
Agree	1. I cared abour Chappie but not that much
Strongly agree	1. the film is able to make the audience feel sympathy for Chappie 2. Chappie was very warm and friendly I was sad at points for them

Qualitative responses for *Chappie* (2015) question 2.4

The film helped increase my understanding about artificial intelligence and consciousness	
Response Category	Comment
Disagree	1. Before watching I already had a pretty good understanding on these topics, but I think the movie asks good questions about those topics. 2. I think it doesnt show a realistic AI

Qualitative responses for *Chappie* (2015) question 2.5

I felt more positive about future artificial intelligence after viewing the film than before viewing it	
Response Category	Comment
Strongly disagree	1. The most realistic aspect of the movie is that AI is very likely going to be used for nefarious purposes in the future
Disagree	1. I think you cant really compare real life and Sci-Fi. Plus we only see the effects of the AI in a very small environment.

Qualitative responses for *Chappie* (2015) question 2.6

I have, at least once, thought about the film when seeing, hearing, or saying something else about artificial intelligence	
Response Category	Comment
Strongly disagree	1. Does not portay AI realisticly 2. Not once I relatade AI with the movie maybe because it is not a very good movie
Disagree	1. here are many other books and films about AI that are far more compelling
Agree	1. When it comes to robot police this Elesium and the new Robocop are strong contenders this portrays them better than either of those two other films

Qualitative responses for *Chappie* (2015) question 2.7

Artificially intelligent humanoid robots will be created that behave in the same way as the robots portrayed in the film	
Response Category	Comment
Disagree	1. I do not believe we will achieve suh levels of conciense
Neither agree nor disagree	1. How in the world would I or anyone taking this survey know this answer? 2. I hope they will 3. There's no way to know that
Agree	1. I find these robots more likely because of the application that is the focus of the film.

Qualitative responses for *Chappie* (2015) question 2.8

I'd like to live in a world in which artificially intelligent humanoid robots, such as those portrayed in the film, exist	
Response Category	Comment
Strongly disagree	1. Humanity hasn't earned the right to create a new species with its own consciousness. Only people with Narcissistic Personality Disorder endeavor to create a new species. Beware the folly of Icarus.
Disagree	1. No, ai is probably the end for humans, but I'd still prefer that people attempt to

	do so, as technological advance is key for human beings 2. I would not. The possibility for misuse is high. 3. Benevolent AI would be fine (even though it's highly unlikely that's possible), AI under the control of corporations is going to be abused
Agree	1. Preferably only the positive aspects apply though. 2. I would love to be surrounded by friendly robots

Qualitative responses for *Chappie* (2015) question 2.9

The film changed my opinion of the possibility of conscious artificially intelligent humanoid robots being created	
Response Category	Comment
Strongly disagree	1. I've always believe it was possible, but I'm not sure whether anyone ever will.
Disagree	1. Again, the movie is ot realistic so I didnt give a clear view of what might happen 2. I think the film is naive and portrays AI like we want to see not like what it will be 3. It will happen eventually.
Neither agree nor disagree	1. Not as much as change my view, but more like harden my opinion on it being possible some day.

Qualitative responses for *Chappie* (2015) question 2.10

I often interact with services or applications which make use of artificial intelligence	
Response Category	Comment
Neither agree nor disagree	1. I don't know how much AI is involved with my searches, map usage, and Amazon purchases.
Agree	1. Now everything uses AI one way or another
Strongly agree	1. any time I use google. 2. Most services online from large providers (facebook, google, etc) that

	deal with user data likely use AI at some point in their service

Qualitative responses for *Chappie* (2015) question 3.1

I know about the trends and use of current and upcoming technology	
Response Category	Comment
Agree	1. Somewhat 2. I keep up 3. Only on the current, not so much on upcoming 4. I keep myself updated regarding this subject and its advances 5. At least those that are public

Qualitative responses for *Chappie* (2015) question 3.2

I am optimistic about future technologies involving Artificial Intelligence	
Response Category	Comment
Disagree	1. Technology is always coopted by the powerful to infringe on people's rights. 2. It's going to be misused and not gonna end well 3. I'm not optimistic about much
Neither agree nor disagree	1. It depends a lot on how much power the corporations that are working 2. AI tech has potential to do a lot of good as well as a lot of bad. Hopefully in time regulations will catch up so this balance weighs in the favor of benefits
Strongly agree	1. I think that they are a tool to make us all better.

Qualitative responses for *Chappie* (2015) question 3.3

Which statement best describes your feelings towards future developments to artificial intelligence?	
Response Category	Comment
I believe that it will ultimately be disastrous/destructive for humans and/or the planet	1. However, none of those disasters would be worse than what someone unethical getting to the technology first would do.

	2. Caveat: only humans, most things that are destructive to humans are fine for the planet
I believe that it will ultimately be bad for humans and/or the planet	1. Humans can't even live with themselves. Why would anyone think they could live with a separate sentient species of artificial origins?
I believe that it will neither be good or bad, and will not significantly change society	1. I think its impossible to know how AI will effect our world considering how little we understand about our own minds let along digital ones 2. I think it will be significant but am not confident as to whether that will be an overall positive or negative
I believe that it will ultimately be good for humans and/or the planet	1. Like I said before. We screw everything up. All I can hope for is a net positive. 2. Depends on the morals of the people implementing it 3. Don't think it's as simple as overall good/bad, it will be like the impact modern technology & the internet has had on society except on a much bigger scale

Qualitative responses for Chappie (2015) question 3.4